

Article

Not peer-reviewed version

The Complete Genome of Two Australian Biotypes of *Fusarium oxysporum* f. sp. *vasinfectum*

[Angel David Popa-Baez](#)*, [Linda J. Smith](#), [Warwick N. Stiller](#), [Melanie Soliveres](#), [Gunjan Pandey](#), [Christopher A. Saski](#), [Don C. Jones](#), [Iain W. Wilson](#)

Posted Date: 13 May 2025

doi: 10.20944/preprints202505.0977.v1

Keywords: *Fusarium* wilt; *Fusarium oxysporum* f. sp. *vasinfectum*; Comparative genomics; Chromosome-level assembly; Australian FOV isolates; Pathogen evolution



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

The Complete Genome of Two Australian Biotypes of *Fusarium oxysporum* f. sp. *vasinfectum*

Angel David Popa-Baez ^{1,*}, Linda J. Smith ², Warwick N. Stiller ³, Melanie Soliveres ¹, Gunjan Pandey ⁴, Christopher A. Sasaki ⁵, Don C. Jones ⁶ and Iain W. Wilson ¹

¹ CSIRO Agriculture and Food, Canberra, ACT 2601, Australia

² Ecosciences Precinct, Department of Primary Industries, Dutton Park, QLD 4102, Australia

³ CSIRO Agriculture and Food, Locked Bag 59, Narrabri, NSW 2390, Australia

⁴ CSIRO Environment, Canberra, ACT 2601, Australia

⁵ Department of Plant and Environmental Sciences, Clemson University, Clemson, South Carolina, USA

⁶ Cotton Incorporated, Cary, North Carolina, USA

* Correspondence: angel.popa@csiro.com

Abstract: *Fusarium oxysporum* f. sp. *vasinfectum* (*Fov*) is a devastating cotton pathogen. Australian *Fov* strains are distinguished by their ability to infect plants without nematode interaction and are genetically distinct from global *Fov*, classified into two vegetative compatibility groups (VCG-01111 and VCG-01112). Here, we present chromosome-level genome assemblies of a historical isolate for each Australian *Fov* VCG. The end-to-end gapless genome assemblies demonstrate high contiguity and completeness, with 97.7% BUSCO completeness for both isolates. Phylogenetic analysis indicates that the Australian *Fov* lineages diverged from other known *Fov* genomes over 3.6 million years ago, while VCG-01111 and VCG-01112 separated approximately 1.1 million years ago. Comparative genomics analysis identified four accessory chromosomes unique to the Australian isolates. Functional annotations revealed 14,495 and 15,342 genes in VCG-01111 and VCG-01112, respectively, with accessory chromosomes containing significantly fewer genes than core chromosomes. Ortholog analysis uncovered unique gene clusters enriched in key metabolic pathways, pathogenicity, and cell division processes. Additionally, we identified several novel lineage-specific peptides unique to each Australian isolates. This comprehensive genomic characterization provides the first insights into the unique evolutionary history of Australian *Fov*, distinguishing them from global *Fov* races. Our findings lay the foundation for understanding the genetic factors underlying their exceptional virulence, which makes Australian *Fov* among the most aggressive cotton pathogens worldwide.

Keywords: fusarium wilt; *Fusarium oxysporum* f. sp. *vasinfectum*; comparative genomics; chromosome-level assembly; Australian FOV isolates; pathogen evolution

1. Introduction

Fusarium wilt (FW) of cotton is caused by the soil-borne fungal pathogen *Fusarium oxysporum* f. sp. *vasinfectum* (*Fov*). The fungus infects cotton roots and colonizes the vascular system, where it proliferates and induces vessel occlusion resulting in plant wilting, stunting, defoliation and death [1,2]. The pathogen's ability to persist in soil for decades as chlamydospores poses a significant challenge for disease management, making the development of resistant cotton cultivars the primary strategy for combating the disease [3–5]. Classical and molecular characterization of *Fov* has led to the identification of multiple races globally [6], (race 1-7), which are classified into 23 recognized and several undetermined vegetative compatibility groups (VCGs) [9]. These VCGs are thought to represent clonal lineages [4,7] with VCG-0111 to VCG-01110 encompassing strains pathogenic to cotton [5]. In culture, Australian *Fov* strains are morphologically identical to each other but are distinguished by the production of unique pigmentation, and a characteristic volatile odour not observed for other *Fov* isolates. Notably, these strains are highly pathogenic in the absence of

nematode infection [2]. Although Australian isolates of *Fov* are similar to race 6 on standard differential hosts assays. Consequently, they have been assigned to distinct vegetative compatibility groups VCGs 01111 and 01112 [2]. Molecular analyses using gene sequences and markers have confirmed the genetic distinctiveness of Australian *Fov* relative to the global *Fov* population. These studies have established that the Australian *Fov* is closely related to a lineage of indigenous *F. oxysporum* found in several areas of Australia, including the Darling Downs in Queensland [8–10].

Australia is indigenous to 17 wild *Gossypium* species, four (*G. australe*, *G. bickii*, *G. nelsonii*, *G. sturtianum*) in which host ranges are known to overlap or abut areas where the majority of cultivated cotton is grown [11]. A survey of *Fusarium* spp. associated with these wild cottons identified a number of *F. oxysporum* isolates that caused mild, but typical FW symptoms on cultivated cotton [10], suggesting that two *Fov* VCGs found in commercial cotton fields evolved from local populations that were associated with native cotton species.

Fusarium wilt (FW) was first identified in Australia in 1993 [12] and spread rapidly, emerging as one of the most economically destructive diseases affecting Australian cotton production. The widespread susceptibility of commercial cultivars resulted in production losses of virtually 100% in some fields [13]. Although the inheritance of host resistance to FW is complex, significant progress has been made in developing cultivars with genetic resistance [14]; however, the disease remains of significant economic concern. The two Australian *Fov* VCGs have different geographic distributions with VCG 01111 occurring in all infected cotton growing regions, while VCG 01112 is restricted to the Boggabilla region of northern NSW where it was first detected. However, cultivar resistance based on vascular browning scores do not appear to be significantly different between the two VCGs [15].

Understanding the evolutionary origin and relationships of pathogens is important for effective disease management. The unique evolutionary trajectory and independent origin of Australian isolates raises important questions about their genomic adaptations and virulence mechanisms, which can be revealed through extensive comparative genomics studies. For instance, gene family evolution and metabolic adaptations in plant pathogens often reflect their environmental pressures and host specialization [16,17]. Understanding these changes can reveal mechanisms of host colonization and tissue specificity, which can help understand deviations and the key characteristics that identify an Australian biotype. For instance, in *F. oxysporum* f. sp. *lycopersici* (*Fol*), the causal agent of tomato wilt, different races present distinct behaviours *in planta*, which can help the pathogen overcome the plant resistance mechanism [18]. In *Fol*, the genome is divided into two groups: core chromosomes and lineage-specific chromosomes. This latter group primarily hosts the pathogenic Secreted in Xylem genes (SIX) [19], highlighting the importance of lineage-specific chromosomes for the pathogenicity of *F. oxysporum* pathogens [20].

In the case of *Fov*, several studies have tried to identify which SIX proteins are shared by the Australian biotypes, which have shown that SIX6 appears to be present in both Australian biotypes but not in other races or VCGs of *Fov*. The limited number of high-quality genomes for the *Fov* races complicates the study of these homologues [21–23]. Although recently, several complete genomes of some *Fov* races have become available. For instance, *Fov* race 7, like in *Fol*, race 7 of *Fov* also contained lineage-specific chromosomes, which also carried ortholog SIX genes like *Fol* and other *Fov* race 7 specific genes [21]. Comparative genomic and functional validation studies have shown that these SIX proteins are crucial factors that contribute to the pathogenicity of *Fov* race 7 [21]. Therefore, studying the evolution of these genomic elements can provide insights into host adaptation and virulence mechanisms.

Thus, given the uniqueness and potential origin of new Australian isolates, we have built two new genome assemblies for the Australian biotype VCG-01111 and VCG-01112 to aid researchers in identifying genetic differences between *Fov* races and potentially uncovering mechanisms capable of increasing cotton resistance to these devastating diseases.

In this study, we generated chromosome-level annotated genome assemblies for two historical Australian isolates of *Fov* (VCG-01111 and VCG-01112) and conducted comparative genomic

analyses to characterise their unique features. We focused on genome architecture, gene family evolution, and pathogenicity determinants, including SIX proteins and novel secreted peptides. This analysis provides insights into the genomic differences between these isolates and other *Fov* strains, revealing potential avenues for understanding their enhanced virulence and regional adaptation.

2. Materials and Methods

Fungi Isolate Collection

Two *Fov* isolates were obtained from the Department of Agriculture and Fisheries (now the Department of Primary Industries), Queensland, Australia culture collection. These isolates were initially recovered from cotton, displaying typical symptoms of Fusarium wilt, grown in fields in Queensland and New South Wales, Australia, during the early to mid-1990s. Specifically, isolate ATCC24233 (VCG-01111) was collected from Cecil Plains in the Darling Downs region of Queensland in May 1993, while isolate ATCC24673 (VCG-01112) was isolated from a property in Boggabilla, MacIntyre Valley, New South Wales in January 1995. The latter strain was isolated from the Siokra 1-4 cotton variety.

DNA Extraction and Quality Assessment

The *Fov* isolates were maintained in 15% glycerol at minus 80°C and recovered in potato dextrose agar (¼ PDA) [24] at 25°C for 4-7 days. Mycelium was aseptically scraped from the plate, snapped, frozen in liquid Nitrogen, and stored at minus 80 °C. Genomic DNA was extracted by The Westmead Institute for Medical Research, using a modified CTAB method [25]. First, 300 mg of dried mycelium powder was mixed with 10 ml CTAB buffer (2% CTAB, 100 mM Tris-HCl pH 8.0, 20 mM EDTA pH 8.0, 1.4 M NaCl) and incubated at 65°C for 30 min. Then, lysate was treated with 50 µg/ml RNase A and 3 M sodium acetate (pH 5.2) at 37°C for 30 min. DNA was purified using sequential extractions with phenol:chloroform:isoamyl alcohol (25:24:1) and chloroform:isoamyl alcohol (24:1), followed by precipitation with an equal volume of isopropanol at -20°C for 2 h. The pellet was washed with 70% and 100% ethanol, air-dried, and dissolved in 0.1 ml TE buffer. The extracted DNA was stored at -80°C until used for HiFi long-read sequencing in a PacBio Revio at the AGRF, Sydney, Australia.

Snap-frozen hyphae from both isolates were submitted for library preparation using a Phase Genomics Proximo Hi-C Fungal kit and sequenced on the Illumina NovaSeq X Series. This work was performed by the Biomolecular Resource Facility – an NCRIS-supported Bioplatforms Australia genomics node at The John Curtin School of Medical Research, ANU.

Genome Assembly and QC

The genome size of both strains was estimated using the default Genomescope2 (v2.0.1) parameters with 19 k-mers [26]. The short (HiC) and long reads (HiFi) were filtered to 60X of the estimated genome size and assembled with default Mabs (hifiasm) under default conditions [27,28]. The resulting assembly was decontaminated with NCBI's FCS tool (v0.5.4) [29] and scaffolded into chromosomes with the default YAHS pipeline (v1.2.1) [30]. The HiC maps were manually curated in Juicebox (v2.20.00) [31] to generate the final genome assembly. Genome assembly quality was assessed using BlobToolKit (v4.2.1), which reviewed the genome assemblies' statistics to produce a Snail plot for each genome. Additionally, the Benchmarking Universal Single-Copy Orthologue (BUSCO) (v5.4.7) score was calculated for each genome, as well as its genome composition.

Genome Annotation

Genome annotation was conducted using the funannotate pipeline (v1.9.1) [32]. First, we remove repetitive contigs with funannotate clean and soft-masked repetitive elements in the remaining scaffolds. Following, we predicted gene sequences using "fusarium" as a BUSCO seed [33]. Functional annotation of the predicted genes was done using interProScan (v5.73-104.0) [34] and antiSMASH (v 7.0) [35] to identify protein families and secondary metabolites. We used SignalP

(version 6) [36] to identify signal peptides using the “Eukarya” as the organism, fast model and short output instructions.

Comparative Genomics and Gene Orthology Analysis

Gene Orthology (GO) and phylogenetic analysis of the assembled and annotated genomes using the OrthoVenn3 comparative genomics platform [37]. For this analysis, we used eight *F. oxysporum* genomes (Table S1), with *Fol* and *F. verticillioides* (*Fve*) serving as outgroups. For the orthologous analysis with OrthoVenn3 (v3), we used the OrthoFinder (v3.0.1) [38] algorithm with an E-value filter of 0.001 and an inflation value of 1.50. The phylogenetic analysis model was JTT+CAT. Gene family expansion and contraction were estimated with a divergence time of 5.34 million years ago (MYA) for *F. oxysporum* and *F. verticillioides*, which was obtained from TimeTree5 [39]; this value was derived from four different studies [40–43]. Finally, we identified a lineage-specific scaffold for the two new genomes by comparing sequence alignments with the *Fol* and *Fve* genome assemblies using D-GENIES (v1.5.0) [44].

3. Results

3.1. Genome Assembly

We generated two highly contiguous and complete de novo genomes for Australian *Fusarium oxysporum* f. sp. *vasinfectum* isolates, VCG-01111 and VCG-01112. The new assemblies contain all 15 chromosomes and showed similar statistics: N50 scaffold length of 4.3 Mb and 4.43 Mb, 47.3% and 47.5% GC content, no gaps, and BUSCO completeness of 97.7% for both FOV VCG-01111 and FOV VCG-01112, respectively (Figure 1). These statistics highlight the contiguity and completeness of these two de novo genome assemblies.

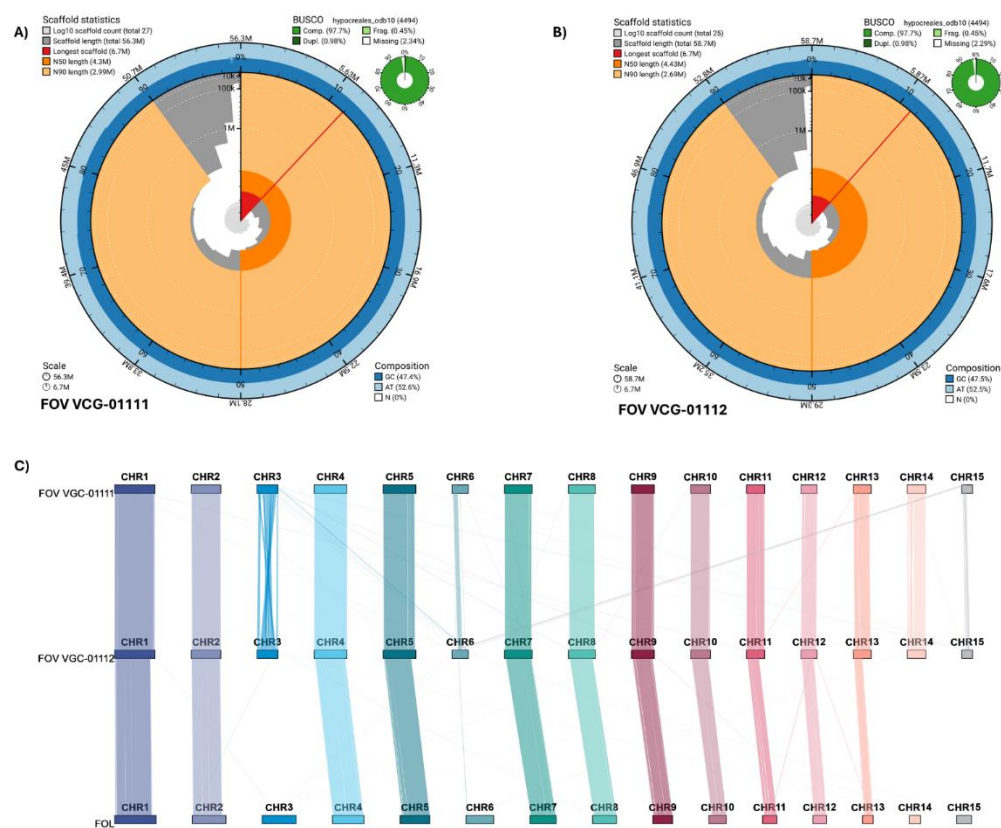


Figure 1. Genome assembly statistics for the two Australian biotypes. (A) Snail plot for the VCG 01111 genome assembly showing the BUSCO and genome contiguity with description of the N50 and N80 length per scaffold, as well as the GC, AT, and N-gaps content; (B) Snail plot showing similar genome statistics for the VCG-01112

genome assembly. (C) Pairwise genome synteny for both Australian biotypes in relations to *Fusarium oxysporum* f. sp. *lycopersicum*.

Pairwise alignment of both assemblies with *F. oxysporum* f. sp. *lycopersici* (Fol) was used to reorient the scaffolds and identify the core and accessory chromosomes. We identified 11 chromosomes that shared a high level of synteny with the Fol chromosomes; thus, we designated them as core chromosomes, and the remaining four (CHR3, CHR6, CHR14, CHR15) were characterised as accessory chromosomes (**Figure 1**).

Three of the accessory chromosomes were not syntenous. First, CHR3 showed a high level of rearrangements within the chromosome, with only the ends of the chromosomes showing synteny between the genomes and a small group of chromosome segments aligning with other core and accessory chromosomes (**Figure 1c**). Second, only partial sections of CHR6 and CHR15 showed collinearity or alignment with other chromosomes, indicating potential regions private to each isolate (**Figure 1c**).

3.2. Genome Annotation and Functional Characteristics

Functional annotation for the VCG-01111 genome identified 14,495 genes, with a total size of 22,333kb and a mean size of 1,483bp, which accounted for 39.57% of the genome. For the VCG-01112, the gene total number was 15,342 with a total size of 22,609 and a mean average of 1,478bpp, which accounted for 38.75% of the genome. We also observed a significantly smaller number of genes in the lineage-specific chromosomes (CHR3, CHR6, CHR14, and CHR15), with only an average of 23% of the VCG-01111 lineage-specific chromosomes containing coding genes and 30% of VCG-01112 lineage-specific chromosomes containing gene coding sequence.

Ortholog analysis for the eight *Fov* genomes revealed 19,414 ortholog clusters, with 6,962 single-copy clusters; this came directly out of the 132,987 proteins present across eight *Fov* genomes, 5,268 (3.96%) of which were singletons. In the case of the Australian genomes, we found 121 singletons for VCG-01111 and 132 for VCG-01112; however, the cluster for both genomes had 442 clusters of orthologs. VCG-01112 had 11 unique gene clusters with 29 genes, whose GO enrichment results showed that the most significant functional specialization was DNA template and transcription (2 genes, $P < 0.001$); this was accompanied by enrichment in the lipid metabolic process (4 genes, $P < 0.001$). In contrast, VCG-01111 only had three gene clusters with seven genes unique to this isolate, which were not enriched.

When we looked at the 442 gene clusters for the Australian genome clusters, GO enrichment analysis revealed significant functional specialization across biological processes, molecular functions, and cellular components. The most significantly enriched biological process was DNA-templated transcription (21 genes, $P < 0.0001$), accompanied by multiple regulatory mechanisms, including both positive and negative regulation of RNA polymerase II-mediated transcription ($P < 0.0001$) (**Table S2**).

Notably, pathogenesis-related processes showed substantial enrichment (9 genes, $P < 0.0001$), indicating potential host-interaction capabilities. This was complemented by enrichment in protein transport (4 genes, $P < 0.0001$) and sporulation processes (4 genes, $P < 0.0001$). Carbohydrate metabolism emerged as a functional category, significantly enriching catabolic and transport processes. Specifically, we observed enrichment in carbohydrate transport (6 genes, $P < 0.0001$) and various polysaccharide degradation pathways, including cellulose (4 genes, $P < 0.0001$), xylan (4 genes, $P < 0.0001$), and pectin (6 genes, $P < 0.0001$) catabolism (**Table S2**).

Additional metabolic capabilities were indicated by enrichment in lipid metabolism ($P < 0.0001$), protein folding ($P < 0.0001$), and various biosynthetic processes, including sterol ($P < 0.0001$) and GPI anchor biosynthesis ($P < 0.0001$). The presence of multiple transmembrane transport processes, including drug transport ($P < 0.0001$) and amino acid transport ($p < 1E-16$), suggests sophisticated metabolite trafficking systems (**Table S2**).

At the cellular component level, significant enrichment was observed in integral membrane components (5 genes, $P < 0.0001$) and nuclear localization (3 genes, $P < 0.0001$), consistent with the observed enrichment in transcription and transport processes. The molecular function analysis revealed a pronounced enrichment in oxidoreductase activity (22 genes, $P < 0.0001$) and zinc ion binding (9 genes, $P < 0.0001$), along with diverse enzymatic activities including transferases (5 genes, $P < 0.0001$), monooxygenases (6 genes, $P < 0.0001$), and protein kinases (4 genes, $P < 0.0001$) (Table S2).

3.3. Evolutionary Differences and Gene Family Evolution

Gene family evolution analysis with CAFE revealed dynamic patterns of gene family expansions and contractions across the phylogenetic timeline spanning approximately 5.3 million years since the separation of *F. oxysporum* and *F. verticillioides*. The analysis identified distinct evolutionary patterns across the different lineages (Figure 2B). The outgroup species *F. verticillioides* showed moderate gene family evolution with 306 expansions and 120 contractions, providing context for this group's ancestral state of gene families (Figure 2A). However, an interesting result is the relationship between the Australian biotypes and the other *Fov* in which *Fol* which had the most dramatic gene family expansion (Figure 2B) is more closely related to all the *Fov* genomes in our study than the two Australian biotypes. In fact, the two Australian biotypes most common recent ancestor with *Fol* and the other *Fov* occurred around ~3.67 MYA preceding the split of *Fol* and the other *Fov* genomes most recent common ancestor. Interestingly, the most recent split in our time-calibrated gene family tree is between Race 4 and Race 7 which occurred around 0.86 MYA. These two races contained the largest gene expansion (+109) of any other node splits. The significant increase in gene family size and the unique set of gene clusters linked to these two races could explain some of their unique pathogenic characteristics, especially in the case of Race 4, which does not need cooperation with nematodes to infect cotton plants. Furthermore, in the same clade we observed that Race 1 occurred ~371,000 years after the split from *Fol* which our phylogenetic analysis estimate occurred ~250,000 years after the most recent common ancestor the preceded the evolution of the two Australian biotypes.

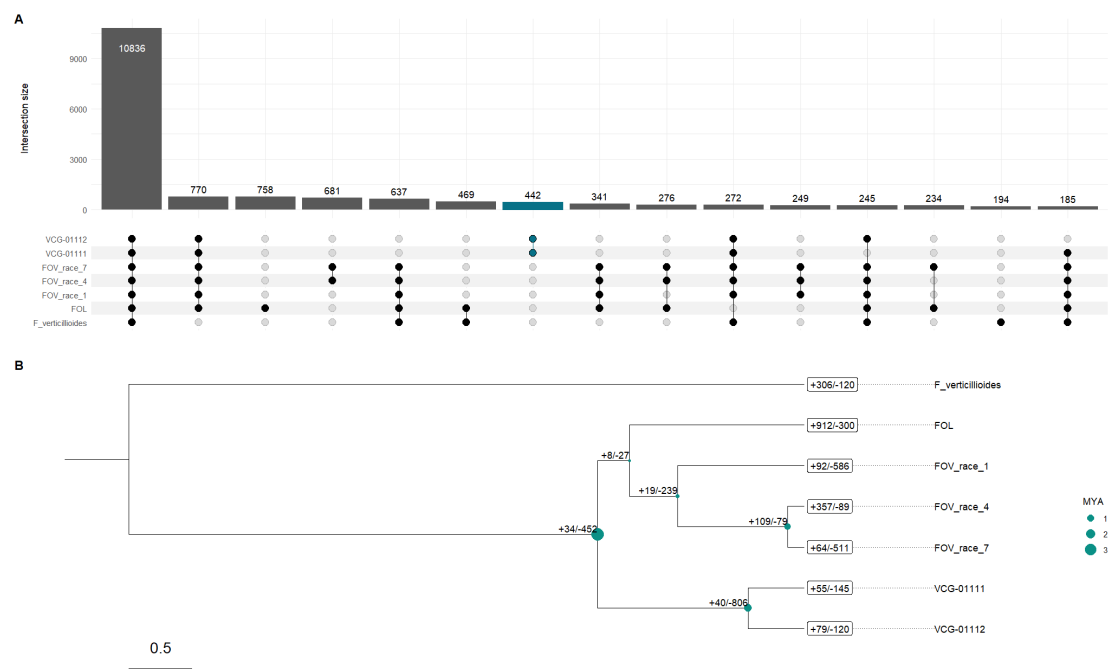


Figure 2. Gene orthology (GO) analysis results and gene family analysis with published *Fov* genomes. (A) Upset plot for the GO clusters found via Orthovenn3 showing the unique gene clusters for different genomes and genome pairs. (B) Phylogenetic tree generated with CAFE5 showing the time calibrated phylogenetic relationship between *Fov* genomes as well as the gene family gain and losses over evolutionary time, branch

length shows times in 0.5 MYA interval calibrated based on the estimated time split between *F. verticillioides* and *F. oxysporum* of 5.3 MYA.

The split between VCG 01111 and VCG 01112 is of particular interest as it revealed contrasting patterns of gene family evolution. VCG-01112 exhibited 79 expanded and 120 contracted gene families, while VCG-01111 showed 55 expansions and 145 contractions (**Figure 2B**). Their immediate common ancestor underwent substantial gene family changes, with 40 expansions and 806 contractions, suggesting a period of significant genome refinement (**Figure 2B**).

Functional enrichment analysis of gene families in VCG-01111 revealed a distinctive pattern of adaptation, particularly in stress response and metabolic processes (**Table S3**). The most significantly enriched biological process was the cellular response to xenobiotic stimulus (31 genes, $P < 0.0001$), complemented by response to antibiotic (19 genes, $P < 0.0001$), and other molecular function involved in biosynthesis and biodegradation of compounds, which suggest sophisticated xenobiotic response mechanisms.

Metabolic functions showed significant specialization, with oxidoreductase activity being highly enriched (9 genes, $P < 0.0001$). Notable metabolic adaptations included polyamine catabolism (18 genes, $P < 0.0001$), styrene metabolism (11 genes, $P < 0.0001$), and alcohol metabolic processes (9 genes, $P < 0.0001$). More specific catabolic processes were also enriched, including tyrosine catabolism (10 genes, $P < 0.0001$), arginine catabolism (4 genes, $P < 0.0001$), and ethanol catabolism (5 genes, $P < 0.0001$) (**Table S3**).

Transcriptional regulation and chromatin modification emerged as important functional categories, with significant enrichment in DNA-templated transcription (15 genes, $P < 0.0001$) and histone acetyltransferase activity (15 genes, $P < 0.0001$). UV-damage excision repair mechanisms (17 genes, $P < 0.0001$) suggest enhanced DNA maintenance capabilities (**Table S3**).

Transport processes showed substantial enrichment, particularly in phosphate ion transmembrane transport (11 genes, $P < 0.0001$) and pantothenate transmembrane transport (11 genes, $P < 0.0001$). Additional transport functions included plasma membrane acetate transport (7 genes, $P < 0.0001$) and riboflavin transport (5 genes, $P < 0.0001$), supported by general transmembrane transporter activity (12 genes, $P < 0.0001$) (**Table S3**).

In the case of VCG-01112, functional enrichment analysis of gene families revealed distinct patterns of functional specialization, particularly in metabolic activities and regulatory processes (**Table S4**). The most significantly enriched molecular function was oxidoreductase activity (9 genes, $P < 0.0001$), complemented by specialized oxidoreductase activities acting on paired donors with oxygen incorporation (38 genes, $P < 0.0001$), suggesting enhanced oxidative metabolism capabilities.

A striking enrichment was observed in nucleotide-binding functions, particularly GTP binding (51 genes, $P < 0.0001$) and GTPase activity (10 genes, $P < 0.0001$), indicating expanded signal transduction and molecular switching capabilities. This was further supported by enrichment in signal transduction pathways (4 genes, $P = 0.014$) and regulatory processes (**Table S4**).

DNA-related processes showed significant enrichment, particularly in UV-damage excision repair (17 genes, $P < 0.0001$), suggesting enhanced DNA maintenance capabilities. The enrichment in histone acetyltransferase activity (15 genes, $P < 0.0001$) and sequence-specific DNA binding transcription factor activity (20 genes, $P < 0.0001$) indicates sophisticated epigenetic and transcriptional regulation mechanisms. This was further evidenced by enrichment in negative regulation of helicase activity (10 genes, $P < 0.0001$) and positive regulation of transcription from RNA polymerase II promoter (3 genes, $p = 0.014$) (**Table S4**).

Pathogenesis-related functions maintained significant representation (11 genes, $P < 0.0001$), along with hyphal growth (5 genes, $P = 0.0015$), suggesting retention of key virulence-associated capabilities. The SOS response pathway (5 genes, $P < 0.0001$) indicates enhanced stress response mechanisms (**Table S4**).

Several metabolic processes showed significant enrichment, including methanol metabolism (10 genes, $P < 0.0001$), sulphur compound metabolism (6 genes, $P < 0.0001$), and terpenoid biosynthesis

(2 genes, $P = 0.0012$). Transport-related processes were also enriched, including riboflavin transport (5 genes, $P < 0.0001$) and plasma membrane acetate transport (7 genes, $P < 0.0001$) (**Table S4**).

Furthermore, the gene family evolution in the common ancestor of VCG-01111 and VCG-01112 revealed significant enrichment patterns associated with both expanded (+40) and contracted (-806) gene families. These families showed substantial enrichment in several key functional categories, particularly those related to stress response and regulatory processes.

Among the most significantly enriched biological processes in the expanded families were stress granule assembly (34 genes, $P < 0.0001$) and transcription elongation from RNA polymerase II promoter (66 genes, $P < 0.0001$), suggesting enhanced capability for transcriptional regulation and stress response. This was further supported by enrichment in regulation of transcription from RNA polymerase II promoter (28 genes, $P < 0.0001$) and regulation of mRNA stability in response to oxidative stress (20 genes, $P < 0.0001$) (**Table S5**).

Notably, pathogenesis-related functions were significantly enriched (20 genes, $P < 0.0001$), along with response to pheromone (24 genes, $P < 0.0001$), indicating retention and expansion of virulence-associated capabilities, which we have already observed in FOV VCG-01112. The expansion also encompassed genes involved in key metabolic processes, including the terpenoid biosynthetic process ($P < 0.0001$) and the phospholipid metabolic process (35 genes, $P < 0.0001$) (**Table S5**).

Regarding molecular functions, the expanded gene families showed significant enrichment in oxidoreductase activity (24 genes, $P < 0.0001$) and metal ion binding (44 genes, $P < 0.0001$). Protein modification capabilities were also enhanced by enrichment in cAMP-dependent protein kinase regulator activity (20 genes, $P < 0.0001$) (**Table S5**).

Signal transduction pathways showed notable modifications, significantly enriching TORC2 signalling (28 genes, $P < 0.0001$), suggesting the refinement of cellular signalling networks, potentially affecting growth regulation and stress responses (**Table S5**).

3.4. Identification of Secreted in Xylem Proteins (SIX)

In addition to identifying gene family evolution, we looked at the SIX proteins from FOL to identify their orthologs in the two Australian genome isolates. Blast analysis revealed the presence of several SIX gene orthologs in both FOV VCG-01111 and FOV VCG-01112. We identified six putative SIX genes (FOXG_04863, FOXG_05755, FOXG_11033, SIX6, FOXG_02829, and FOXG_04805) through sequence homology searches. The most significant hits were found in both vegetative compatibility groups, with multiple high-scoring sequence alignments suggesting strong conservation of these effector proteins. This comparison noted that apart from SIX6, no other lineage-specific SIX was identified in the Australian genomes since all other SIX genes (FOXG_04863, FOXG_05755, FOXG_11033, FOXG_02829, and FOXG_04805) are non-lineage-specific SIX. Interestingly, we observed two genes with high sequence similarity to SIX6 in FOV VCG-01111 and FOV VCG-01112 (**Table S6**).

As noted earlier, we observed 121 singletons or unassigned genes for VCG-01111 and 132 for VCG-01112 and a few hundred genes specific to both lineages. Thus, there is potential that some of the species-specific genes are also potential signal peptides that contribute to the pathogenesis of our two genome assemblies. As such, we predicted signal peptides in all our protein sequences. We identified 1,228 genes from the FOV VCG-01111 genome assembly, while FOV VCG-01112 had 1,265 genes classified as signal peptides. After placing the signal peptides, we double-checked that the ortholog signal peptides identified earlier were also identified as signal peptides in our analysis. **Table S7** shows that our analysis classified all the SIX genes orthologs as signal peptides. Given this validation, we looked at the potential for the unassigned genes to be putative lineage-specific single peptides, and we found that out of the 121 unassigned genes for FOV VCG-01111, five (4.12%) were putative signal peptides. Meanwhile, FOV VCG-01112 had only 3 (2.27%) of the 132 unassigned genes as putative signal peptides (**Table S8**). Thus, the remaining species-specific unassigned ortholog genes may be involved in functions contributing to these isolates' pathogenicity and adaptation.

4. Discussion

In this study, we have constructed two highly contiguous and annotated genomes for the Australian *Fov* biotypes VCG-01111 and VCG-01112. These genomes were assembled using a third-generation sequencing approach that incorporates long reads (PacBio HiFi) and 3D chromosome structural information (HiC) resulting in two complete genome assemblies. These two new genomes join the existing list of 12 *Fov* genomes currently available at NCBI [45] (as of February 2025), providing a valuable addition to the *Fov* genome resources.

Our phylogenetic analyses reveal a significant evolutionary distance between the Australian biotypes of *Fov* and other global biotypes, corroborating previous findings that link VCG-01111 and VCG-01112 more closely to native Australian wild *Fusarium* species than to other *Fov* races [10]. These results reinforce the hypothesis that VCG-01111 and VCG-01112 are native to Australia and may have undergone a host switch following the introduction of cultivated cotton. Supporting this idea, initial identification of VCG-01111 occurred in cotton fields surrounded by native Australian *Gossypium* species [46]. Our estimates suggest that the evolutionary divergence between Australian *Fov* strains and other *Fov* occurred approximately 3.4 MYA (**Figure 2b**), predating the polyploidization of cultivated cotton (1–2 MYA) but following the divergence of native Australian *Gossypium* (5–10 MYA) [47].

A key phenotypic distinction between Australian *Fov* strains and most other *Fov* strains is the absence of nematode interaction as a prerequisite for pathogenic infection of cotton [48]. While the devastating Race 4 strain is also capable of infecting cotton in the absence of nematodes, Australian biotypes have not been detected outside of Australia, suggesting an independent evolutionary origin of this trait. Given that Race 4 is closely related to Race 7 and Race 1 [21,22] and has undergone a more recent evolutionary split (**Figure 2b**), future studies should investigate the underlying molecular mechanisms that enable *Fov* strains, including Race 4, VCG-01111, and VCG-01112, to infect cotton without nematode interaction.

Our comparative genomic analyses did not reveal significant clustering between the Australian biotypes and Race 4 at the gene level. While both Australian biotypes share portions of the lineage-specific chromosomes, they do not share all lineage-specific elements. Given that lineage-specific chromosomes are known to play a role in *Fusarium* pathogenicity and host adaptation, it is expected that genetic differences influencing pathogenicity would reside within these regions. However, we did not observe a significant enrichment of signal peptides or gene numbers relative to chromosome size in the lineage-specific chromosomes, contrasting with findings from the Race 7 *Fov* genome, where lineage-specific chromosomes harbour genes contributing to pathogenicity [21]. Nonetheless, gene families associated with pathogenicity are enriched in the Australian genomes, and lineage-specific chromosomes do contain other signal peptides that could contribute to adaptation and virulence.

Notably, genes unique to both Australian isolates are associated with extensive transcriptional regulation, secretory pathways, complex polysaccharide metabolism, and metabolite transport systems. This suggests that the Australian genomes have evolved enhanced transcriptional control, diverse metabolic capabilities, particularly in carbohydrate utilization and sophisticated transport mechanisms. These traits, combined with the enrichment of pathogenesis-related genes and hydrolytic enzymes, indicate potential adaptation of the Australian biotypes for cotton infection. This is particularly relevant given the regional specificity of at least one of these genomes [10,49,50] and observed field phenotypic variations from standard *Fov* disease presentations [8]. However, genetic marker analyses suggest that these differences are unlikely to be due to the emergence of new or undetected isolates [9,51]. The observed genomic complexity supports the notion that variations in disease symptoms may arise from differential activation of intricate cellular pathways within the pathogen.

We identified orthologs of secreted-in-xylem (SIX) proteins, which are crucial for fungal pathogenicity, in the Australian biotype genomes through BLAST analysis. However, this only included the six putative SIX genes, including SIX6, a known ortholog found in the Australian

biotypes, which differentiated them from other *Fov* isolates lacking this gene [50]. The other five orthologs identified are not considered lineage-specific proteins for *Fol*; however, they do appear to be differentially expressed across different stages of infection development, as well as act as novel candidate effector genes that cause cell death. [53]. The initial characterisation of these effector genes gives insight into their potential mechanism orchestrated by the pathogen during early infection stages. For instance, FOXG_0405 in *Fol* is significantly upregulated during the first two days post-inoculation. In comparison, FOXG_11033 is predominantly upregulated within four days post-inoculation, and the remaining four effector genes (FOXG_02829, FOXG_04863, FOXG_05755, SIX6) were most up-regulated at six days post-infection [53]. This orchestration of gene effectors at the early stages of infection could point towards distinct effects these proteins have in the plant cells required for successful invasion. Thus, the presence of these effectors in the Australian biotype genomes could indicate the likely molecular mechanism implemented by these pathogens to infect the cotton plants. To our knowledge, this is the first time these proteins have been identified within the Australian biotypes of *Fov*, which highlights the value of whole genome sequencing, annotation and evolutionary characterisation of pathogenic species.

While these SIX orthologs are likely contributors to pathogenicity, they may not be the sole determinants. We identified 121 genes without homologs, some of which encode predicted signal peptides. These findings open new avenues for proteomics and functional genomic studies to elucidate the infection mechanisms employed by Australian biotypes, which may differ from those of other *Fov* isolates.

Given the high similarity in host range and virulence among Australian biotypes [46], genetic differences beyond pathogenicity, are likely associated with adaptation to environmental conditions. VCG-01111 is widely distributed, while VCG-01112 remains geographically restricted to its initial discovery region. One theory for this difference is that VCG-01111 was inadvertently spread via large earth moving equipment that was situated in the Darling Downs where it was discovered. The Come Clean Go Clean procedures initiated after the discovery of *Fov* then restricted further spread of *Fov*, preventing movement of VCG-01112. However, considering the time elapsed since the initial discovery of *Fov* in Australia and the extensive cultivation of cotton, it's plausible that VCG-01111 possesses greater environmental adaptability than VCG-01112. This is corroborated by recent field data showing that cotton fields in Boggabilla, in New South Wales, were traditionally infected by VCG-01112 are now dominated by VCG-01111, indicating a complete displacement of VCG-01112 (Linda J. Smith per personal communication). The differences between the VCG-01111 and VCG-01112 genomes do not suggest a facile answer to their current distribution, but the lower gene count in VCG-01111 may contribute to its broader adaptability, while VCG-01112 may be more suited to a specialized ecological niche. Gene family evolution analyses indicate a complex and dynamic evolutionary trajectory, with notable genomic differences between the two Australian isolates despite their close genetic relationship. These differences may be driven by distinct ecological niches and structural genome variations [54,55].

We also observed that many functionally enriched genes involved cellular processes related to enhanced xenobiotic resistance and stress tolerance, supported by diverse metabolic capabilities and sophisticated transport systems. The combination of enriched drug response genes, stress signaling pathways, and various catabolic processes indicates the evolution of robust survival mechanisms, potentially contributing to environmental resilience and adaptation in the Australian *Fov* isolates. These patterns of gene family evolution suggest multiple instances of independent adaption and given the presence of other cotton species in Australia, this pathogen may have continued to acquire and lose functions that aid with host specificity, virulence, and metabolic capabilities.

5. Conclusions

In this study, we have developed two new genome assemblies for Australian *Fov* biotypes VCG-01111 and VCG-01112. Through comparative genomic analysis, we identified key genetic and functional differences that distinguish these biotypes from other *Fov* strains. These new genome

assemblies provide a valuable resource for understanding *Fov* pathogenicity, resistance mechanisms, and the evolutionary processes that shaped these unique Australian biotypes.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org, Table S1: Genome assemblies used for genomic comparisons and phylogenetic analysis. Table S2: Gene Orthology Enrichment Results for the 442 Australian isolates unique gene clusters; Table S3: Gene Orthology Enrichment Results VCG-01111 gene family expansion and contraction; Table S4: Gene Orthology Enrichment Results VCG-01112 gene family expansion and contraction; Table S5: Gene Orthology Enrichment Results for the common-ancestor of VCG-01111 and VCG-01112 gene family expansion and contraction; Table S6: Results from reciprocal best hit analysis of both VCG-01111 and VCG-01112 genes against the SIX gene from *Fov* race 7; Table S7: Signal peptide classification for all SIX genes orthologs in the VCG-01111 and VCG-01112 genome; Table S8: Novel putative signal peptides identified for each of the genomes.

Author Contributions: I.W.W, C.A.S, D.C.J and W.N.S. conceived the study. M.S. and L.J.S. performed the experiments. A.D.P-B. and G.P. analysed the data. A.D.P-B, I.W.W., W.N.S and L.J.S, wrote the manuscript. C.A.S and D.C.J. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Cotton Breeding Australia (a joint venture between Cotton Seed Distributors Ltd. and CSIRO) Cotton Incorporated, Cotton Research and Development Corporation and the Foundation for Food and Agriculture Research (Grant #23-000853).

Data Availability Statement: Raw reads with adapters removed and genome assemblies with basic gene annotation are available at NCBI with the accession number under accessions SAMN45866337 and SAMN45866338. Genome assemblies with annotations are deposited under the BioProject PRJNA1199160, accession numbers JBMJCH000000000 and JBMJCI000000000 for VCG-01111 and VCG-01112 respectively. The genome assemblies can also be found in GenBank under GenBank accession ids GCA_049307005.1 and GCA_049306905.1 for both VCG-01111 and VCG-01112 respectively. We have also deposited the genomes and the annotation on the CSIRO Data Access Portal [56] with permalink: <https://doi.org/10.25919/13pj-dn18>.

Acknowledgments: The authors would like to thank David Zheng of The Westmead Institute for Medical Research for his work on isolating high-quality DNA from our *Fov* isolates for HiFi long-read sequencing, and Maxim Nekrasov from the Biomolecular Resource Facility at the Australian National University for his Hi-C sample preparation and sequencing. We would also like to acknowledge the added support from CSIRO Global.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Davis, R.M.; Colyer, P.D.; Rothrock, C.S.; Kochman, J.K. Fusarium Wilt of Cotton: Population Diversity and Implications for Management. *Plant Disease* **2006**, *90*, 692–703, doi:10.1094/PD-90-0692.
2. Davis, R.D.; Moore, N.Y.; Kochman, J.K. Characterisation of a Population of Fusarium Oxysporum f.Sp. Vasinfectum Causing Wilt of Cotton in Australia. *Australian journal of agricultural research* **1996**, *47*, 1143, doi:10.1071/ar9961143.
3. Chakrabarti, A. Fusarium Oxysporum: A “Moving” View of Pathogenicity. In *Soil Biology*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2013; pp. 157–189 ISBN 978-3-642-39338-9.
4. Bentley, S.; Kochman, J.K.; Moore, N.Y.; Pattemore, J.A.; Gulino, L.; O'Neill, W.T. DNA Diagnostics for Fusarium Wilt of Cotton.; Australian Cotton Growers' Research Association, January 1 2000; pp. 455–461.
5. Molecular Characterization of Races and Vegetative Compatibility Groups in Fusarium Oxysporum f. Sp. Vasinfectum | Applied and Environmental Microbiology Available online: <https://journals.asm.org/doi/abs/10.1128/aem.60.11.4039-4046.1994> (accessed on 13 February 2025).
6. Kim, Y.; Hutmacher, R.B.; Davis, R.M. Characterization of California Isolates of Fusarium Oxysporum f. Sp. Vasinfectum. *Plant disease* **2005**, *89*, 366–372, doi:10.1094/pd-89-0366.

7. Fernandez D; Assigbese K; Dubois M P; Geiger J P Molecular Characterization of Races and Vegetative Compatibility Groups in *Fusarium Oxysporum* f. Sp. *Vasinfectorum*. *Applied and Environmental Microbiology* **1994**, *60*, 4039–4046, doi:10.1128/aem.60.11.4039-4046.1994.
8. Le, D.P.; Tran, T.T.; Gregson, A.; Jackson, R. TEF1 Sequence-Based Diversity of *Fusarium* Species Recovered from Collar Rot Diseased Cotton Seedlings in New South Wales, Australia. *Australasian Plant Pathol.* **2020**, *49*, 277–284, doi:10.1007/s13313-020-00706-8.
9. Le, D.P.; Nguyen, C.P.T.; Kafle, D.; Scheikowski, L.; Montgomery, J.; Lambeth, E.; Thomas, A.; O’Keeffe, K.; Shakeshaft, B.; Young, A.; et al. Surveillance, Diversity and Vegetative Compatibility Groups of *Fusarium Oxysporum* f. Sp. *Vasinfectorum* Collected in Cotton Fields in Australia (2017 to 2022). *Pathogens* **2022**, *11*, 1537, doi:10.3390/pathogens11121537.
10. Wang, B.; Brubaker, C.L.; Summerell, B.A.; Thrall, P.H.; Burdon, J.J. Local Origin of Two Vegetative Compatibility Groups of *Fusarium Oxysporum* f. Sp. *Vasinfectorum* in Australia: Evolutionary Origin of Cotton Wilt in Australia. *Evolutionary applications* **2010**, *3*, 505–524, doi:10.1111/j.1752-4571.2010.00139.x.
11. Craven, L.; Stewart, J.; Brown, A.; Grace, J. The Australian Wild Species of *Gossypium*.; January 1995.
12. Kochman, J. *Fusarium* Wilt in Cotton — a New Record in Australia. *Australasian Plant Pathology* **1995**, *24*, 74–74, doi:10.1071/APP9950074.
13. Stiller, W.N.; Wilson, I.W. Australian Cotton Germplasm Resources. In *World Cotton Germplasm Resources*; IntechOpen, 2014 ISBN 978-953-51-1622-6.
14. Conaty, W.C.; Broughton, K.J.; Egan, L.M.; Li, X.; Li, Z.; Liu, S.; Llewellyn, D.J.; MacMillan, C.P.; Moncuquet, P.; Rolland, V.; et al. Cotton Breeding in Australia: Meeting the Challenges of the 21st Century. *Front. Plant Sci.* **2022**, *13*, doi:10.3389/fpls.2022.904131.
15. McFadden, H.; Beasley, D.; Brubaker, C.L. Assessment of *Gossypium Sturtianum* and *G. Australe* as Potential Sources of *Fusarium* Wilt Resistance to Cotton. *Euphytica* **2004**, *138*, 61–72, doi:10.1023/B:EUPH.0000047076.38747.81.
16. Baroncelli, R.; Amby, D.B.; Zapparata, A.; Sarrocco, S.; Vannacci, G.; Le Floch, G.; Harrison, R.J.; Holub, E.; Sukno, S.A.; Sreenivasaprasad, S.; et al. Gene Family Expansions and Contractions Are Associated with Host Range in Plant Pathogens of the Genus *Colletotrichum*. *BMC Genomics* **2016**, *17*, 555, doi:10.1186/s12864-016-2917-6.
17. Krijger, J.-J.; Thon, M.R.; Deising, H.B.; Wirsal, S.G. Compositions of Fungal Secretomes Indicate a Greater Impact of Phylogenetic History than Lifestyle Adaptation. *BMC Genomics* **2014**, *15*, 722, doi:10.1186/1471-2164-15-722.
18. Swett, C.L.; Del Castillo Múnera, J.; Hellman, E.; Helpio, E.; Gastelum, M.; Lopez Raymundo, E.; Johnson, H.; Oguchi, R.; Hopkins, A.; Beaulieu, J.; et al. Monitoring for a New I3 Resistance Gene-Breaking Race of *F. Oxysporum* f. Sp. *Lycopersici* (*Fusarium* Wilt) in California Processing Tomatoes Following Recent Widespread Adoption of Resistant (F3) Cultivars: Challenges with Race 3 and 4 Differentiation Methods. *Frontiers in plant science* **2023**, *14*, 1088044, doi:10.3389/fpls.2023.1088044.
19. Zhang, Y.; Yang, H.; Turra, D.; Zhou, S.; Ayhan, D.H.; DeIulio, G.A.; Guo, L.; Broz, K.; Wiederhold, N.; Coleman, J.J.; et al. The Genome of Opportunistic Fungal Pathogen *Fusarium Oxysporum* Carries a Unique Set of Lineage-Specific Chromosomes. *Commun Biol* **2020**, *3*, 1–12, doi:10.1038/s42003-020-0770-2.
20. Yu, D.S.; Outram, M.A.; Smith, A.; McCombe, C.L.; Khambalkar, P.B.; Rima, S.A.; Sun, X.; Ma, L.; Ericsson, D.J.; Jones, D.A.; et al. The Structural Repertoire of *Fusarium Oxysporum* f. Sp. *Lycopersici* Effectors Revealed by Experimental and Computational Studies. *eLife* **2024**, doi:10.7554/elife.89280.2.
21. Yang, D.; Zhang, X.; Ming, Y.; Liu, C.; Zhang, X.; Liu, S.; Zhu, L. Characterization of the High-Quality Genome Sequence and Virulence Factors of *Fusarium Oxysporum* f. Sp. *Vasinfectorum* Race 7. *Journal of fungi (Basel, Switzerland)* **2024**, *10*, 242, doi:10.3390/jof10040242.
22. Jobe, T.O.; Ulloa, M.; Ellis, M.L. Two de Novo Genome Assemblies from Pathogenic *Fusarium Oxysporum* f. Sp. *Vasinfectorum* Race 4 (FOV4) Isolates from California. *Microbiology resource announcements* **2024**, *13*, e0076023, doi:10.1128/mra.00760-23.
23. Jobe, T.O.; Ulloa, M.; Ellis, M.L. A High-Quality Whole-Genome Sequence, Assembly, and Gene Annotation of *Fusarium Oxysporum* f. Sp. *Vasinfectorum* (Fov) Race 1 from California. *Microbiology resource announcements* **2024**, *13*, e0070223, doi:10.1128/mra.00702-23.

24. Nash, S.M.; Snyder, W.C. Quantitative Estimations by Plate Counts of Propagules of the Bean Root Rot *Fusarium* in Field Soils. *Phytopathology* **1962**, *52*, 567–572.
25. Murray, M.G.; Thompson, W.F. Rapid Isolation of High Molecular Weight Plant DNA. *Nucleic Acids Research* **1980**, *8*, 4321–4326, doi:10.1093/nar/8.19.4321.
26. Vurture, G.W.; Sedlazeck, F.J.; Nattestad, M.; Underwood, C.J.; Fang, H.; Gurtowski, J.; Schatz, M.C. GenomeScope: Fast Reference-Free Genome Profiling from Short Reads. *Bioinformatics* **2017**, *33*, 2202–2204, doi:10.1093/bioinformatics/btx153.
27. Schelkunov, M.I. Mabs, a Suite of Tools for Gene-Informed Genome Assembly. *BMC Bioinformatics* **2023**, *24*, 377, doi:10.1186/s12859-023-05499-3.
28. Cheng, H.; Concepcion, G.T.; Feng, X.; Zhang, H.; Li, H. Haplotype-Resolved de Novo Assembly Using Phased Assembly Graphs with Hifiasm. *Nat Methods* **2021**, *18*, 170–175, doi:10.1038/s41592-020-01056-5.
29. Rapid and Sensitive Detection of Genome Contamination at Scale with FCS-GX - PMC Available online: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10898089/> (accessed on 18 February 2025).
30. Zhou, C.; McCarthy, S.A.; Durbin, R. YaHS: Yet Another Hi-C Scaffolding Tool. *Bioinformatics* **2023**, *39*, btac808, doi:10.1093/bioinformatics/btac808.
31. Durand, N.C.; Robinson, J.T.; Shamim, M.S.; Machol, I.; Mesirov, J.P.; Lander, E.S.; Aiden, E.L. Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst* **2016**, *3*, 99–101, doi:10.1016/j.cels.2015.07.012.
32. Palmer, J.M.; Stajich, J. Funannotate v1.8.1: Eukaryotic Genome Annotation 2020.
33. Simão, F.A.; Waterhouse, R.M.; Ioannidis, P.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO: Assessing Genome Assembly and Annotation Completeness with Single-Copy Orthologs. *Bioinformatics* **2015**, *31*, 3210–3212, doi:10.1093/bioinformatics/btv351.
34. InterProScan 5: Genome-Scale Protein Function Classification | Bioinformatics | Oxford Academic Available online: <https://academic.oup.com/bioinformatics/article/30/9/1236/237988?login=true> (accessed on 18 February 2025).
35. Blin, K.; Shaw, S.; Augustijn, H.E.; Reitz, Z.L.; Biermann, F.; Alanjary, M.; Fetter, A.; Terlouw, B.R.; Metcalf, W.W.; Helfrich, E.J.N.; et al. antiSMASH 7.0: New and Improved Predictions for Detection, Regulation, Chemical Structures and Visualisation. *Nucleic Acids Research* **2023**, *51*, W46–W50, doi:10.1093/nar/gkad344.
36. SignalP 6.0 Predicts All Five Types of Signal Peptides Using Protein Language Models | Nature Biotechnology Available online: https://www.nature.com/articles/s41587-021-01156-3?utm_campaign=related_content&utm_source=HEALTH&utm_medium=communities (accessed on 18 February 2025).
37. Sun, J.; Lu, F.; Luo, Y.; Bie, L.; Xu, L.; Wang, Y. OrthoVenn3: An Integrated Platform for Exploring and Visualizing Orthologous Data across Genomes. *Nucleic Acids Research* **2023**, *51*, W397–W403, doi:10.1093/nar/gkad313.
38. OrthoFinder: Phylogenetic Orthology Inference for Comparative Genomics | Genome Biology | Full Text Available online: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-019-1832-y> (accessed on 18 February 2025).
39. Kumar, S.; Suleski, M.; Craig, J.M.; Kaspruwicz, A.E.; Sanderford, M.; Li, M.; Stecher, G.; Hedges, S.B. TimeTree 5: An Expanded Resource for Species Divergence Times. *Molecular Biology and Evolution* **2022**, *39*, msac174, doi:10.1093/molbev/msac174.
40. Shen, X.-X.; Steenwyk, J.L.; LaBella, A.L.; Opulente, D.A.; Zhou, X.; Kominek, J.; Li, Y.; Groenewald, M.; Hittinger, C.T.; Rokas, A. Genome-Scale Phylogeny and Contrasting Modes of Genome Evolution in the Fungal Phylum Ascomycota. *Sci Adv* **2020**, *6*, eabd0079, doi:10.1126/sciadv.abd0079.
41. Da Lage, J.-L.; Binder, M.; Hua-Van, A.; Janeček, S.; Casane, D. Gene Make-up: Rapid and Massive Intron Gains after Horizontal Transfer of a Bacterial α -Amylase Gene to Basidiomycetes. *BMC Evol Biol* **2013**, *13*, 40, doi:10.1186/1471-2148-13-40.
42. Van der Nest, M.A.; Steenkamp, E.T.; McTaggart, A.R.; Trollip, C.; Godlonton, T.; Sauerman, E.; Roodt, D.; Naidoo, K.; Coetzee, M.P.A.; Wilken, P.M.; et al. Saprophytic and Pathogenic Fungi in the Ceratocystidaceae Differ in Their Ability to Metabolize Plant-Derived Sucrose. *BMC Evol Biol* **2015**, *15*, 273, doi:10.1186/s12862-015-0550-7.

43. O'Donnell, K.; Rooney, A.P.; Proctor, R.H.; Brown, D.W.; McCormick, S.P.; Ward, T.J.; Frandsen, R.J.N.; Lysøe, E.; Rehner, S.A.; Aoki, T.; et al. Phylogenetic Analyses of RPB1 and RPB2 Support a Middle Cretaceous Origin for a Clade Comprising All Agriculturally and Medically Important Fusaria. *Fungal Genet Biol* **2013**, *52*, 20–31, doi:10.1016/j.fgb.2012.12.004.
44. Cabanettes, F.; Klopp, C. D-GENIES: Dot Plot Large Genomes in an Interactive, Efficient and Simple Way. *PeerJ* **2018**, *6*, e4958, doi:10.7717/peerj.4958.
45. Sayers, E.W.; Bolton, E.E.; Brister, J.R.; Canese, K.; Chan, J.; Comeau, D.C.; Connor, R.; Funk, K.; Kelly, C.; Kim, S.; et al. Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **2022**, *50*, D20–D26, doi:10.1093/nar/gkab1112.
46. Davis, R.D.; Moore, N.Y.; Kochman, J.K. Characterisation of a Population of *Fusarium Oxysporum* f.Sp. *Vasinfestum* Causing Wilt of Cotton in Australia. *Aust. J. Agric. Res.* **1996**, *47*, 1143–1156, doi:10.1071/ar9961143.
47. Taxonomy and Evolution of the Cotton Genus, *Gossypium* - Wendel - 2015 - Agronomy Monographs - Wiley Online Library Available online: <https://access.onlinelibrary.wiley.com/doi/abs/10.2134/agronmonogr57.2013.0020> (accessed on 21 March 2025).
48. Wagner, T.A.; Duke, S.E.; Davie, S.M.; Magill, C.; Liu, J. Interaction of *Fusarium* Wilt Race 4 with Root-Knot Nematode Increases Disease Severity in Cotton. *Plant Disease* **2022**, doi:10.1094/PDIS-12-21-2725-SC.
49. Moore, N.Y.; Kochman, J.K.; Obst, N.R.; O'Neill, W.T.; Bentley, S. *Fusarium* Wilt of Cotton in Australia.
50. Chakrabarti, A.; Rep, M.; Wang, B.; Ashton, A.; Dodds, P.; Ellis, J. Variation in Potential Effector Genes Distinguishing Australian and Non-Australian Isolates of the Cotton Wilt Pathogen *Fusarium Oxysporum* f.Sp. *Vasinfestum*. *Plant Pathology* **2011**, *60*, 232–243, doi:10.1111/j.1365-3059.2010.02363.x.
51. Genetic Variation and Population Structure of *Fusarium Oxysporum* f.Sp. *Vasinfestum* in Australia - Wang - 2006 - Plant Pathology - Wiley Online Library Available online: <https://bsppjournals.onlinelibrary.wiley.com/doi/full/10.1111/j.1365-3059.2006.01445.x> (accessed on 18 February 2025).
52. Disease Ranks. Cotton Seed Distributors.
53. Sun, X.; Fang, X.; Wang, D.; Jones, D.A.; Ma, L. Transcriptome Analysis of *Fusarium*-Tomato Interaction Based on an Updated Genome Annotation of *Fusarium Oxysporum* f. Sp. *Lycopersici* Identifies Novel Effector Candidates That Suppress or Induce Cell Death in *Nicotiana Benthiana*. *Journal of fungi (Basel, Switzerland)* **2022**, *8*, 672, doi:10.3390/jof8070672.
54. R. D. Davis, N. Y. Moore, J. K. Kochman Characterisation of Population of *Fusarium Oxysporum* f. Sp. *Vasinfestum* Causing Wilt of Cotton in Australia. *Aust. J. Agric. Res* **1996**, *47*, 1143–1156.
55. Lopez-Lavalle, L.A.B.; Gillespie, V.J.; Tate, W.A.; Ellis, M.H.; Stiller, W.N.; Llewellyn, D.L.; Wilson, I.W. Molecular Mapping of a New Source of *Fusarium* Wilt Resistance in Tetraploid Cotton (*Gossypium Hirsutum* L.). *Mol Breeding* **2012**, *30*, 1181–1191, doi:10.1007/s11032-012-9705-z.
56. Popa-Baez, A.; Smith, L.J.; Stiller, W.; Soliveres, M.; Pandey, G.; Wilson, I. The Complete Genome of Two Australian Isolates of *Fusarium Oxysporum* f. Sp. *Vasinfestum*. **2025**, doi:10.25919/13pj-dn18.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.