

candidate_genes_discovery

lihui xiang

2025-04-02

I. Literature Reference

Otero et al. (2024) – *Evaluation of the Genetic Diversity, Haplotype, and Virulence of Fusarium oxysporum f. sp. vasinfectum* Field Isolates from Alabama (<https://apsjournals.apsnet.org/doi/10.1094/PHYTO-11-23-0438-R>)

This study revealed that the race 4 isolate 89-1A carries a set of small accessory chromosomes, including those harboring lactamase-related genes. These accessory chromosomes are postulated to contribute to the virulence of *Fusarium oxysporum* f. sp. *vasinfectum* on cotton, suggesting a potential role of horizontally transferred elements in pathogenicity.

II.Conclusion from the reference

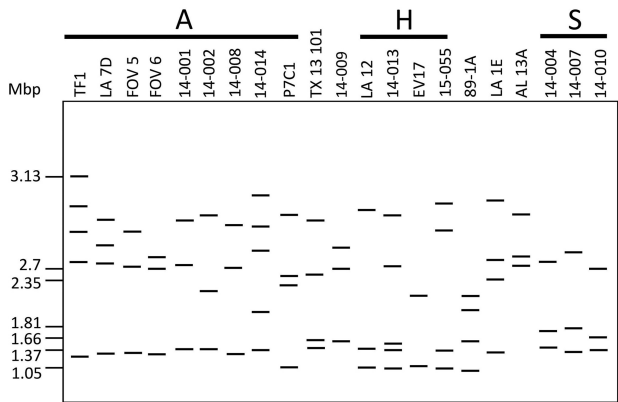


Figure 1. Small chromosome diversity in 21 *Fusarium oxysporum* f. sp. *vasinfectum* (Fov) field isolates assessed by pulsed-field gel electrophoresis.(Otero et al. 2024)

The race 4 isolate **89-1A** shows distinct small chromosome bands compared to other isolates. These unique chromosomes are likely accessory chromosomes and may contain virulence-associated genes, such as lactamases, contributing to increased pathogenicity on cotton.

III.Localizing the Genes of Interest from a Specific Accessory Chromosome

Table 1: Location of the potential cotton-specific gene cluster
(Otero et al. 2024)

Domain, (Sense)	89-1A Protein	Length	CA-12 Locus
Cupin, signal peptide (-)	FUN_000082, VEDA_05193	1,330 bp, 371 AA	1-1,330
Glutathione S-transferase (-)	FUN_000109, VdG1_00948	755 bp, 233 AA	1,975- 2,729
Uncharacterized protein (-)	FUN_000146	447 bp, 148 AA	4,978- 5,424
GMC oxidoreductase, signal peptide (-)	FUN_000208, VEDA_05195	2,118 bp, 621 AA	10,007- 12,124
Uncharacterized protein (+)	FUN_000330	821 bp, 248 AA	19,825- 20,644
SAM-dependent methyltransferase (-)	FUN_000376	1,109 bp, 211 AA	22,395- 23,503
SAM-dependent methyltransferase (-)	FUN_000387	1,392 bp, 225 AA	23,531- 24,922
FMN-dependent dehydrogenase (+)	FUN_000469	998 bp, 278 AA	28,966- 29,963
zinc-binding alcohol dehydrogenase (-)	FUN_000504	309 bp, 102 AA	31,458- 31,766
Alpha fold: possible integral membrane protein (+)	FUN_000631	375 bp, 101 AA	42,479- 42,853
Alpha fold: possible integral membrane protein (+)	FUN_000642	746 bp, 230 AA	42,936- 43,681
SAM-dependent methyltransferase (-)	FUN_000920	848 bp, 226 AA	74,477- 75,324
Uncharacterized protein, signal peptide (-)	FUN_000947	848 bp, 135 AA	77,718- 78,565
Uncharacterized protein (-)	FUN_000956	2,316 bp, 771 AA	79,888- 82,203
Uncharacterized protein (-)	FUN_000985	219 bp, 72 AA	83,682- 83,900
Alpha fold: F-box domain-containing protein (-)	FUN_001024	414 bp, 137 AA	88,219- 88,632
Uncharacterized protein, secreted peptide (-)	FUN_001040	890 bp, 257 AA	90,190- 91,079
SEST-like, SGNH hydrolase subfamily, plant suberin lipid hydrolase (+)	FUN_001051	4,914 bp, 1401 AA	91,772- 96,685
AB hydrolase superfamily (+)	FUN_001144	857 bp, 245 AA	102,581- 103,437
Uncharacterized protein (+)	FUN_001170	219 bp, 72 AA	105,029- 108,112
Uncharacterized protein (+)	FUN_001185	2,316 bp, 771 AA	109,590- 111,905
MFS superfamily (+)	FUN_001224	1,260 bp, 419 AA	113,926- 115,185
short-chain dehydrogenase	FUN_001275	1,427 bp, 284 AA	119,041- 120,467
Alpha fold: Nitrogen regulatory protein areA GATA-like domain-containing protein, LisH domain-containing protein (+)	FUN_001318	624 bp, 170 AA	123,450- 124,073

Domain, (Sense)	89-1A Protein	Length	CA-12 Locus
Animal heme peroxidase-like superfamily (+)	FUN_001354	1,713 bp, 447 AA	127,416- 129,128
Animal heme peroxidase-like superfamily, cytochrome P450 superfamily (+)	FUN_001373	2,078 bp, 568 AA	129,362- 131,439
FAD-binding, GlcD superfamily	FUN_001409	1,719 bp, 572 AA	133,828- 135,546
AB hydrolase, signal peptide (+)	FUN_001427	1,293 bp, 430 AA	136,208- 137,499
alpha/beta hydrolase (+)	FUN_001452	1,663 bp, 452 AA	139,638- 141,300
HeLo, HET prion, signal peptide (+)	FUN_001479	1,302 bp, 288 AA	142,100- 143,401
Alpha/beta hydrolase superfamily (-)	FUN_001518	1,333 bp, 248 AA	147,119- 148,451
Integrase, HTH Tc5 transposase DNA-binding protein (+)	FUN_019948	1,844 bp, 551 AA	152,051- 153,894
Lactamase, signal peptide (-)	FUN_001580, VEDA_05199	1,333 bp, 416 AA	155,120- 156,453
Zinc independent alcohol dehydrogenase (-)	FUN_001597, VEDA_05198	1,088 bp, 678 AA	157,732- 158,820
Short-chain dehydrogenase (-)	FUN_001610, VEDA_05197	1,975 bp, 346 AA	158,821- 160,796
MFS aflatoxin efflux pump (+)	FUN_001624, VEDA_05198	3,227 bp, 570 AA	160,937- 164,164
GAL4-like zinc DNA-binding, sterol uptake (+)	FUN_001654	1,175 bp, 364 AA	164,353- 165,528
SNU13 (-)	FUN_001849	1,054 bp, 217 AA	189,259- 190,313
Ankyrin repeats (-)	FUN_001881	449 bp, 149 AA	193,530- 193,979
sulfatase, alkaline phosphatase (-)	FUN_001892	2,258 bp, 733 AA	195,037- 197,296

In **Table 1**, we identified a cluster of genes located on a specific accessory chromosome of *Fusarium oxysporum* f. sp. *vasinfectum* strain CA-12. These genes are absent in most other *F. oxysporum* strains and show high similarity to a conserved region in *Verticillium dahliae* strain Vd991.

Among the listed genes, we focused on **eight candidates** of particular interest. This selection was based on the following rationale:

1. Seven of these genes were previously reported as homologs shared between *F. oxysporum* and *V. dahliae* Vd991.
2. Our current analysis identified an **additional homolog**, expanding the set of conserved genes to eight (Chen et al., 2018).
3. These genes include **cupin**, **GMC oxidoreductase**, **GST**, **FMN-dependent dehydrogenase**, and multiple **SAM-dependent methyltransferases**, all of which are potentially involved in **host-specific adaptation**.
4. The presence of secretion signals and domain predictions in several genes supports their possible role as **accessory virulence factors**.
5. The entire cluster resides on a region absent from core chromosomes, suggesting its horizontal origin and role in **cotton-specific pathogenicity**.

This accessory cluster may therefore serve as a valuable genomic signature for Fov race 4 and a candidate target for disease resistance breeding in cotton.

VI. Localizing the Genes of Interest from a Specific Accessory Chromosome

Based on the previous analysis, we selected **eight genes of interest** located on a putative accessory chromosome for further investigation. These genes were chosen due to their predicted functions, presence of signal peptides, and high similarity to virulence-related genes in *Verticillium dahliae*. For details, see **Table 2** below.

Table 2: Candidate Genes list from the Accessory Chromosome

Domain, (Sense)	89-1A Protein	Length	CA-12 Locus
Cupin, signal peptide (-)	FUN_000082, VEDA_05193	1,330 bp, 371 AA	1-1,330
Glutathione S-transferase (-)	FUN_000109, VdG1_00948	755 bp, 233 AA	1,975-2,729
GMC oxidoreductase, signal peptide (-)	FUN_000208, VEDA_05195	2,118 bp, 621 AA	10,007-12,124
Lactamase, signal peptide (-)	FUN_001580, VEDA_05199	1,333 bp, 416 AA	155,120- 156,453
Zinc independent alcohol dehydrogenase (-)	FUN_001597, VEDA_05198	1,088 bp, 678 AA	157,732- 158,820
Short-chain dehydrogenase (-)	FUN_001610, VEDA_05197	1,975 bp, 346 AA	158,821- 160,796
MFS aflatoxin efflux pump (+)	FUN_001624, VEDA_05198	3,227 bp, 570 AA	160,937- 164,164
GAL4-like zinc DNA-binding, sterol uptake (+)	FUN_001654	1,175 bp, 364 AA	164,353- 165,528