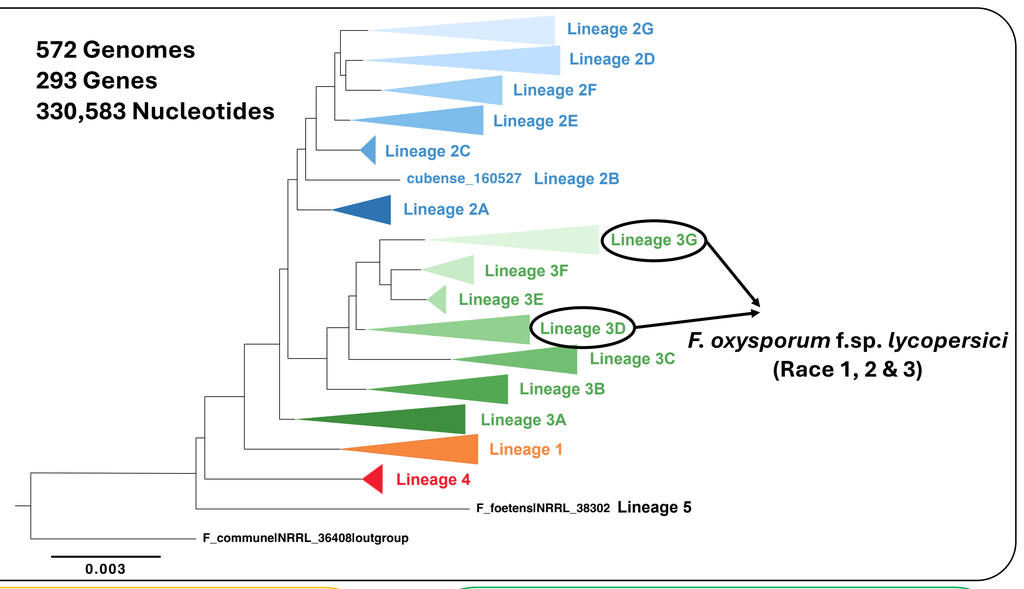
Development of TaqMan qPCR Diagnostic Markers Using E-Probes

Saleh Rahimlou[1], Frank N. Martin[2], Andres S. Espindola[3], Daniel Ramos Lopez[3], David M. Geiser[1]
[1]Pennsylvania State University , [2]USDA Salinas, [3]Oklahoma State University

Fusarium oxysporum possesses a complex evolutionary structure and a highly dynamic genome[1,2], complicating the development of reliable molecular diagnostics for distinguishing pathogenic races. In this study, we focused on developing diagnostic assays for the three major races of the tomato wilt pathogen *F. oxysporum* f. sp. *lycopersici* (Fol). A key challenge is that each of these races occurs in two distinct evolutionary lineages (3D and 3G), requiring markers that are race-specific rather than lineage-specific. To address this, we developed a genomics-informed workflow and Python-based pipeline to identify DNA regions uniquely associated with a target race and suitable for amplification-based detection platforms such as TaqMan qPCR. The approach uses short race-specific sequences (e-probes), identified with the MiFi platform[3,4], which are arranged as triplets within amplification-compatible genomic regions. In each triplet, one e-probe defines the forward primer site, a second defines the internal fluorescent probe region, and a third defines the reverse primer site, collectively forming a candidate diagnostic amplicon. While the objective of this project is development of diagnostic markers for the tomato wilt pathogen, the bioinformatic approach can also be used for marker design for other fungi for use in isolate identification and population studies.



1 Candidate Amplicon Discovery

Reference genome

BLASTn-short
Probe Alignment

Probe 1 Probe 2 Probe 3

Adjacent Probes
(within span limits)

Extract Candidate Amplicon

Probe 1 Probe 2 Probe 3

...ATCGATCGATCGATCGATCG...

2 Genome Specificity Test

All Generated Amplicons

Non-target Genomes

BLAST Each Probe
Against Non-target Genomes

Evaluate Off-target Hits

Probe 1
Probe 2
Probe 3

Classification

PASS (No significant off-target hits)

PARTIAL PASS (Limited off-target hits)

FAIL (Significant off-target hits)

3 Raw Read Specificity Test

Unique e-Probe candidates
(from Step 2)

Raw Sequencing Reads
(FASTQ)

Bowtie2 Mapping

Probe	Total Reads	Mapped Reads
	------	--------
	------	--------
	------	--------

Pass amplicons with no non-target e-probe matches.

REPORT

4 Inclusivity Test

Candidate Amplicons

Target Genomes

Extract Required Probes
(from each amplicon)

Probe 1
Probe 2
Probe 3

BLAST Against All Target Genomes

	Genome 1	Genome 2	Genome 3
Probe 1	✓	✓	✓
Probe 2	✓	✗	✓
Probe 3	✓	✓	✗

Pass if 100% of Target Genomes
Contain All Probes

Inclusive Amplicons

Module 1 – Candidate Amplicon Discovery

Purpose: Identify genomic regions that contain multiple nearby e-probes and define candidate diagnostic amplicons.
What happens:
- Existing e-probes are aligned to the reference genome using BLAST.
- Clusters of neighboring probes are identified based on user-defined spacing.
- The genomic region spanning these probes is extracted as a candidate amplicon for downstream validation.

Output: Candidate amplicon sequences.

Module 2 – Genome Specificity Assessment

Purpose: Remove candidate amplicons that produce e-probe matches in non-target genomes.
What happens:
- Candidate e-probes are searched against a database of non-target genomes.
- Probe matches are evaluated to identify potential cross-reactivity.
- Candidate amplicons are classified as **Pass**, **Partial Pass**, or **Fail** based on off-target hits.

Output: Genome-specific amplicon candidates.

Module 3 – Raw Read Validation

Purpose: Evaluate e-probe performance directly on sequencing reads.
What happens:
- Sequencing reads are aligned to validated e-probes using Bowtie2.
- Read alignments are quantified to confirm probe detection.
- Detection statistics are summarized to assess real-world specificity and robustness.

Output: Final candidate diagnostic amplicons for which all three e-probes show no matches to non-target genomes.

Module 4 – Inclusivity Test

Purpose: Ensure that each specific amplicon is consistently detected across all target isolates.
What happens:
- Specific amplicons are screened against all target genomes.
- Each required e-probe is evaluated for successful detection.
- Only amplicons detected in every target genome are retained.

Output: Highly inclusive candidate diagnostic amplicons.

Model Case: Diagnostic Assay Development for *Fusarium oxysporum* f. sp. *lycopersici*
Objective: Develop race-specific diagnostic amplicons for Fol Races 1, 2, and 3, despite each race occurring in two independent evolutionary lineages.

Pipeline Module	Race 3 Example (CS65 vs. Near Neighbors)
Candidate e-probes	6,040
Candidate amplicons (130–650 bp)	2,956
Genome-specific amplicons	2,091
Raw read-validated amplicons	994–1,204
Inclusivity	21/53 target genomes

Acknowledgements: This research was supported by the USDA National Institute of Food and Agriculture (NIFA) through the Agriculture and Food Research Initiative (AFRI) (Grant No. 500000031486), The Pennsylvania State Experiment Station Project 4837, USDA Agricultural Research Service (ARS) Project 2038-22000-016-00D, and the California Department of Food and Agriculture through the California Specialty Crop Block Grant Program (Grant Nos. 18-0001-059-SC and 21-0001-050-SF).

Visit My Website:
https://rahimlou.github.io

1. Geiser, D. M., Al-Hatmi, A. M., Aoki, T., Arie, T., Balmas, V., Barnes, I., ... & Viljoen, A. (2021). Phylogenomic analysis of a 55.1-kb 19-gene dataset resolves a monophyletic Fusarium that includes the Fusarium solani species complex. *Phytopathology®*, *111*(7), 1064-1079.
2. O'Donnell, K., Ward, T. J., Robert, V. A., Crous, P. W., Geiser, D. M., & Kang, S. (2015). DNA sequence-based identification of Fusarium: current status and future directions. *Phytoparasitica*, *43*(5), 583-595.
3. Stobbe, A. H., Daniels, J., Espindola, A. S., Verma, R., Melcher, U., Ochoa-Corona, F., ... & Schneider, W. (2013). E-probe Diagnostic Nucleic acid Analysis (EDNA): a theoretical approach for handling of next generation sequencing data for diagnostics. *Journal of microbiological methods*, *94*(3), 356-366.
4. Espindola, A. S., & Cardwell, K. F. (2021). Microbe Finder (MiFi®): Implementation of an interactive pathogen detection tool in metagenomic sequence data. *Plants*, *10*(2), 250.