Transposable Elements in Fusarium oxysporum & Growth Inhibition of Fusarium oxysporum Using Pepper Extracts

Taylor Aguiar

Follow this and additional works at: https://scholarworks.umass.edu/masters_theses_2

Part of the Agriculture Commons, Bioinformatics Commons, Computational Biology Commons, and the Integrative Biology Commons

Recommended Citation

This Open Access Thesis is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Masters Theses by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
Transposable Elements in *Fusarium oxysporum* & Growth Inhibition of *Fusarium oxysporum* using Pepper Extracts

A Thesis Presented

by

TAYLOR JAMES AGUIAR

Submitted to the Graduate School of the University of Massachusetts Amherst in the partial fulfillment of the requirements for the degree of

Master of Science

May 2018

Plant Biology
Transposable Elements in *Fusarium oxysporum* & Growth Inhibition of *Fusarium oxysporum* using Pepper Extracts

A Thesis Presented

By

TAYLOR AGUIAR

Approved as to style and content by:

Li-Jun Ma, Chair

Ludmila Tyler, Member

Sibongile Mafu, Member

Li-Jun Ma, Graduate Program Director of Plant Biology
DEDICATION

To my passed friend Tyler Ayan.
Wherever you are, I hope your favorite plants are there.
ACKNOWLEDGEMENTS

To start off, I would like to thank my adviser, mentor, and teacher, Dr. Li-Jun Ma, who has financially supported me throughout my years as both an undergraduate and graduate researcher. I always appreciated all suggestions I received to improve my project and go further with it. Thank you to Dr. Sibongile Mafu for her expertise and patience in teaching me how to use the HPLC. A thank you is also due to Dr. Ludmila Tyler who initially got me very interested in plant biology and helped me through the initial stages of writing.

A massive thank you to my graduate student mentor, Dilay Hazal Ayhan, who patiently taught me everything related to computational biology. Her guidance and knowledge of coding were invaluable throughout the project. I want to thank all my friends and graduate students who were always willing to help improve my thesis throughout the writing process along with any troubleshooting of experiments including Matt James, Kathryn Vescio, and Greg Delulio.

My sincere gratitude goes out to Tracy James Powell who gave me the original peppers, from which I got the seeds. Without him, I never would have experienced the pain of hot peppers and considered the effect of extracts on fungi. I want to thank Chris Joyner and David O’Neil from the CNS greenhouse, who took care of my pepper plants during the later stages.

I want to thank Pei-lun for initially teaching me the screening procedure. I extend my gratitude to Yong Zhang who was always around to help me with the plate reader if it was not functioning properly. I want to express my appreciation to Dr. Li Gou for the
RNA-seq data used in this project. In addition, I want to thank Christina Lopez Díaz, Antonio di Pietro, and Dilay Hazal Ayhan for the genomic data of Fo4287.

A special thank you to my family and friends who were always there to support me and help me stay focused on the goal. The encouragement I received got me through the long days and late nights.
ABSTRACT

TRANSPOSABLE ELEMENTS IN *Fusarium oxysporum* & GROWTH INHIBITION OF *Fusarium oxysporum* USING PEPPER EXTRACTS

MAY 2018

TAYLOR JAMES AGUIAR B.S. UNIVERSITY OF MASSACHUSETTS AMHERST
M.S. PLANT BIOLOGY UNIVERSITY OF MASSACHUSETTS AMHERST
Directed by: Professor Li-Jun Ma

The following contains two projects focused on the fungal pathogen, *Fusarium oxysporum*. The first project was purely computational in the examination of transposable elements (TEs), which are mobile sequences with the ability to multiply and move in their host genome. In *F. oxysporum*, TEs such as miniature impala elements are associated with the secreted in xylem gene that are related to its virulence over its host. The *F. oxysporum* species complex can be utilized as a model system for the examination of TE content and TE expression during the infection cycle. To find whether TEs play a role in the infection process and if their expression changes when fungi are in planta, a comparison was made using RNA-seq data from a pathogenic (*Fo*5176) and a non-pathogenic strain (*Fo*47) of *F. oxysporum* interacting with the model plant *Arabidopsis thaliana*. Complementary to this, the copy numbers of the same TEs were calculated in the two aforementioned strains and in *F. oxysporum* f.sp. *lycopersici* 4287 (*Fo*4287) to find if there was a correlation between expression and copy number. Using these two different datasets together showed that TE expression and copy number are lower in the non-pathogenic strain and unlinked in the infection course.

The second project examined the growth inhibition of *Fusarium oxysporum* isolates *Fo*32931 (the isolate pathogenic to immunocompromised humans) and *Fo*4287 with the use of extracts from chilies of *Capsicum chinense*. Pepper plants were grown
from seed and the peppers were harvested for an ethanol (100%) extraction. After preparation, the optical density of growth of the *F. oxysporum* isolates was measured for a 48-hour period with 96-well plate containing varying concentrations of the extracts and controls. Growth curves were analyzed and normalized to a growth control. After doing High Performance Liquid Chromatography, an estimated concentration of capsaicin (the causal agent of the burning sensation from hot chilis) was established. A correlation between the amount of growth inhibition and the concentration of capsaicin was made. Taken together, the data suggests that an increase of capsaicin concentration in extracts is correlated with reduced growth for the two tested isolates of *F. oxysporum*. 
CONTENTS

ACKNOWLEDGEMENTS ................................................................................................. iv
ABSTRACT .................................................................................................................. vi
LIST OF TABLES .......................................................................................................... x
LIST OF FIGURES ........................................................................................................ xi

CHAPTER
1. TRANSPOSABLE ELEMENTS .................................................................................. 1

Prelude ......................................................................................................................... 1

Introduction ................................................................................................................ 4

Classification: .......................................................................................................... 5
Functions of TEs in Their Host: ............................................................................... 7
Exaptation of TEs and Genome Organization: .................................................. 10

Methods .................................................................................................................... 12

Clean-up and Validation: ....................................................................................... 14

Results & Discussion............................................................................................... 15

Clustering Based on Copy Number ................................................................. 15
Correlation Between Copy Number and Expression .................................. 18

2. PEPPER EXTRACTS AND GROWTH INHIBITION OF F. OXYPORUM .... 25

Introduction ................................................................................................................ 25

Methods ..................................................................................................................... 26

Initial Screens and Set-up .................................................................................... 26
Growing Chili Peppers......................................................................................... 27
Extraction of Capsaicin ....................................................................................... 29
<table>
<thead>
<tr>
<th>Chapter/Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Performance Liquid Chromatography</td>
<td>30</td>
</tr>
<tr>
<td>Optical Density Screens with Extracts</td>
<td>32</td>
</tr>
<tr>
<td>Computational Analysis</td>
<td>32</td>
</tr>
<tr>
<td>Results &amp; Discussion</td>
<td>33</td>
</tr>
<tr>
<td>HPLC Results</td>
<td>33</td>
</tr>
<tr>
<td>Optical Density Screens with Extracts</td>
<td>37</td>
</tr>
<tr>
<td>3. SUPPLEMENTARY FIGURES</td>
<td>46</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>58</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1 Contains the observed patterns repressed expression of all tested TEs for Fo5176. For each pattern, the TE name in bold is the TE shown in the respective plot. The TE expression level that does not involve the interaction with the host is shown with a “c” for control. Plots have an x-axis in hours (post inoculation) and y-axis normalized expression. All expression plots are in Figure 14. Error bars indicate the standard deviation of the samples.................................................................20

Table 2 Contains the observed patterns induced expression of all tested TEs for Fo5176. For each pattern, the TE name in bold is the TE shown in the respective plot. The TE expression level that does not involve the interaction with the host is shown with a “c” for control. Plots have an x-axis in hours (post inoculation) and y-axis normalized expression. All expression plots are in Figure 14. Error bars indicate the standard deviation of the samples.................................................................21

Table 3 This table refers to the chromatograms of Figure 4. These are the areas calculated by auto-integration for each of the peaks at ~6.6 minutes from the capsaicin standard serial dilution. Capsaicin concentrations were corrected for actual capsaicin in the 95% standard. Area is measured in milli-absorption units (mAU) multiplied by time. Exact retention times are also shown in the left-most column. Area values and capsaicin concentrations shown here are represented in Figure 5 in a calibration curve. .........................................................33

Table 4 This table refers to the chromatograms of 50ex (figure 6) and chromatograms of unfiltered extracts (Figure 16). Listed here are the areas calculated by auto-integration for each of the peaks at ~6.6 minutes. Area is measured in milli-absorption units (mAU) multiplied by time. Exact retention times are also shown in the third column. The estimated amounts of capsaicin for each of the samples are listed in the right most column in mg/mL. Each sample was in 1 mL of 100% ethanol.................................................................37
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Plot of all 50 TEs after computational workflow. Copy number resides on the top half while the expression heatmap of Fo5176 and Fo47 lies on the bottom half. Clusters are separated by magenta vertical lines and based on pattern similarity. The copy number plot’s vertical axis is in logarithmic scale. The expression heatmap’s vertical axis is in hours post inoculation with the range of the heatmap going from 0 to 15 (normalized expression values).</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>Plot of Fo5176 TE expression at 0 hpi and genomic copy number.</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>Plot of Fo47 TE expression at 0 hpi and genomic copy number.</td>
<td>19</td>
</tr>
<tr>
<td>4</td>
<td>Overlay of all 11 chromatograms from the 95% capsaicin standard. Y-axis is measured in milli-absorption units (mAU) and x-axis is measured in minutes retention time. The resulting chromatogram are from a 1:2 serial dilution with 95% capsaicin and ethanol. Corrected capsaicin concentrations on Table 4. Capsaicin is responsible for peak at ~6.6 min retention time. Peak was present at ~6.6 min for all but lowest dilution. Areas from the 11 peaks at ~6.6 min were used for standard curve (Figure 5).</td>
<td>33</td>
</tr>
<tr>
<td>5</td>
<td>Calibration curve of capsaicin serial dilution. Exact area values are shown in table 4. Black line connects each of the serial dilution area values. Dotted red line is linear line of best fit. Maximum capsaicin standard was 1 mg/mL capsaicin (0.95 mg/mL actual). Equation for line of best fit is on right side of the figure with an R² value of 0.9983.</td>
<td>34</td>
</tr>
<tr>
<td>6</td>
<td>Y-axis is measured in milli-absorption units (mAU) and x-axis is measured in minutes retention time. Extract sample is listed on the left. WC = Habanero White Chili, CR= Carolina Reaper Chili, MSC = Moruga Scorpion Chili, CRC = Carolina Reaper Chili. These chromatograms are from the 50ex corresponding extract. Arrows designates capsaicin associated peak. As shown by chromatograms of standard capsaicin, (figure 4) capsaicin is responsible for peak at ~6.6 min retention time.</td>
<td>36</td>
</tr>
</tbody>
</table>
Growth inhibition of Fo32931 for each of the pepper extracts, nystatin, and 95% capsaicin standard (x-axis). Areas under the curve for each extract and control (Figure 20) were normalized by the area under the curve of growth in ethanol and converted to percentage form (y-axis). A t-test was done for each treatment in comparison to ethanol growth to find significance of growth inhibition. All significant values are shown with an asterisk (p<0.05). The spore concentration of each well (with culture) was ~1.26E6 spores/mL. Plate was incubated at 25 °C. and was covered.

Growth inhibition of Fo4287 for each of the pepper extracts, nystatin, and 95% capsaicin standard (x-axis). Areas under the curve for each extract and control (Figure 21) were normalized by the area under the curve of growth in ethanol and converted to percentage form (y-axis). A t-test was done for each treatment in comparison to ethanol growth to find significance of growth inhibition. All significant values are shown with an asterisk (p<0.05). The spore concentration of each well (with culture) was ~1.25E6 spores/mL. Plate was incubated at 25 °C. and was covered.

Correlation plot between growth inhibition (%) values of Fo32931 (shown with x points) and Fo4287 (shown with o points) and the corresponding amount of capsaicin in 50ex pepper extracts (shown by table 5). All samples were corrected for capsaicin that was present in the well.

Plot showing growth curves for the 48 hours screen against Fo32931, with a focus on the two MSC extracts. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25 °C and a spore concentration of ~1.26E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested.
(absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested.

12 Plot showing growth curves for the 48 hours screen against Fo32931, with a focus on growth with 0.25 mg/mL capsaicin. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25 °C and a spore concentration of ~1.26E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested.

13 Plot showing growth curves for the 48 hours screen against Fo4287, with a focus on growth with 0.25 mg/mL capsaicin. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25 °C and a spore concentration of ~1.25E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested.

14 Overview figure with 50 subplots showing the expression pattern of each mapped TE for Fo5176 during the infection course. The vertical axis of each subplot changes to stay within the range of the normalized expression. X-axis is measured in hours post inoculation. Plots with a single point means the isolate did not have the TE or mapped expression reads across all time points. TEs that had similar patterns are listed in table 1.

15 Overview figure with 50 subplots showing the expression pattern of each mapped TE for Fo47 during the infection course. The vertical axis of each subplot changes to stay within the range of the normalized expression. X-axis is measured in hours post inoculation. Plots with a single point means the isolate did not have the TE or mapped expression reads across all time points. TEs that had similar patterns are listed in table 2.

16 Vertical axis bar is not consistent across all samples. Y-axis is measured in milli-absorption units (mAU) and x-axis is measured in minutes retention time. Extract sample is listed on the left. WC = Habanero White Chili, CR = Carolina Reaper Chili, MSC = Moruga Scorpion Chili, CRC = Carolina Reaper Chili. These
chromatograms are from 1mL of unfiltered (u) liquid extract from the designated cultivar that was evaporated. Before being run in the HPLC, 1mL of ethanol (100%) was added and the solution mixed. Arrows designates capsaicin associated peak. As shown by chromatograms of standard capsaicin, (figure 4) capsaicin is responsible for peak at ~6.6 min retention time.

This plot is an extension of the linear line of best fit referring to the capsaicin standard and the associated calibration curve. Black line connects each of the serial dilution area values. Dotted red line is linear line of best fit. Maximum capsaicin standard was 1 mg/mL capsaicin (0.95 mg/mL actual). Equation for line of best fit is on right side of the figure with an R2 value of 0.9983.

Subplots showing each of the growth curves for the 48 hours screen against Fo32931. The spore concentration of each well (with culture) was ~1.26E6 spores/mL. Absorbance was measure at 600 nm. Plate was incubated at 25 °C. and was covered. All plots have an x-axis measured in time (n = 30minutes) and a y-axis with OD (absorbance). Each subplot has their condition labeled above their plot. Each line represents one replicate. Controls reside in the three bottom right panels.

Subplots showing each of the growth curves for the 48 hours screen against Fo4287. The spore concentration of each well (with culture) was ~1.25E6 spores/mL. Absorbance was measured at 600 nm. Plate was incubated at 25 °C. and was covered. All plots have an x-axis measured in time (n = 30minutes) and a y-axis with OD (absorbance). Each subplot has their condition labeled above their plot. Each line represents one replicate. Controls reside in the three bottom right panels.

Overview for the 48 hours screen against Fo32931. Bar graph showing the total area (y-axis) under each of the growth curves when in the presence of each pepper extract or control (x-axis). The spore concentration of each well (with culture) was ~1.26E6 spores/mL. Plate was incubated at 25 °C. and was covered. Controls: Normal growth, ethanol, and Nystatin are the left-most bars.
Overview for the 48 hours screen against Fo4287. Bar graph showing the total area (y-axis) under each of the growth curves when in the presence of each pepper extract or control (x-axis). The spore concentration of each well (with culture) was ~1.25E6 spores/mL. Plate was incubated at 25 °C, and was covered. Controls: Normal growth, ethanol, and Nystatin are the left-most bars. 

Plot representing the effect of trying to do an optical density reading with wells containing capsaicin. Spore concentration in each respective well was ~1.8E6 spores/mL. The x-axis is time in hours and the y-axis is a measure of the absorbance. 

Plot showing growth curves for the 48 hours screen against Fo32931, with a focus on the two WC extracts. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25°C and a spore concentration of ~1.26E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested. 

Plot showing growth curves for the 48 hours screen against Fo4287, with a focus on the two WC extracts. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25°C and a spore concentration of ~1.25E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested. 

Plot showing growth curves for the 48 hours screen against Fo32931, with a focus on the two CR extracts. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25°C and a spore concentration of ~1.26E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested. 

Plot showing growth curves for the 48 hours screen against Fo4287, with a focus on the two CR extracts. The color key for each line is in the top left of the figure. Screen conditions included: an
incubation temperature of 25°C and a spore concentration of ~1.25E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested. .................................................. 55

27 Plot showing growth curves for the 48 hours screen against Fo32931, with a focus on the two CRC extracts. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25°C and a spore concentration of ~1.26E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested. .................................................. 55

28 Plot showing growth curves for the 48 hours screen against Fo4287, with a focus on the two CRC extracts. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25°C and a spore concentration of ~1.25E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested. .................................................. 56

29 This figure represents the line of best fit for the correlation between the replicates of growth inhibition for Fo32931 with each of the extracts (shown in Figure 9 with “x”) and the corresponding concentration of capsaicin. ........................................................................... 56

30 This figure represents the line of best fit for the correlation between the replicates of growth inhibition for Fo4287 with each of the extracts (shown in Figure 9 with “x”) and the corresponding concentration of capsaicin. ........................................................................... 57
CHAPTER 1
TRANSPOSABLE ELEMENTS

Prelude

The interaction between plants and pathogenic microbes through evolutionary time can be compared to an “arms race” (Dawkins & Krebs, 1979). On a microscopic level, the “armaments” of the attacker - the pathogen - are usually in the form of secreted effector proteins, which allow the pathogen to evade the defenses of the host cell. The defenses in this arms race metaphor are immune receptors which when triggered, can result in a variety of responses including programmed cell death to prevent the spread of infection (Toruño, Stergiopoulos, & Coaker, 2016). If a specific effector of a pathogen can cause infection, then it is a “virulence factor”. Over time, once the pathogen’s host has gained a new immune receptor (through evolution) that can recognize the effector, the effector then becomes an “avirulence factor” (e.g. AVR gene). The host gene responsible for the immune receptor is therefore a resistance (R) gene (Stahl & Bishop, 2000). Simply put, the “arms race” is a cycling of selection between new virulence factors and new R genes that result in recognition of the effector.

The evolution of the interaction between plants and their pathogens (and the respective R genes related to the failed effector, now AVR, genes) has been a periodic competition long before organized agriculture. In an unforeseen way, the spread of large scale agriculture has had detrimental consequences due to a lack of biodiversity of agriculturally important crops.

For a large-scale and consistent agricultural system to be efficient and potentially mechanized, growth cycles must be homogenous across the individual plants of a single
crop. With this idea, modern agriculture has put a heavy dependence on the use of the monoculture system - a crop of individuals with identical genetics. The use of clonal plants also prevents high variation in phenotypes such as height, which increases the ease with which a crop can be harvested mechanically.

In nature, the arms race is balanced - neither the host nor the pathogen is forced into extinction. In a classic monoculture system, identical plants are used in every growing season, so the host plant is unable to evolve resistance against a pathogen. If one individual in the crop becomes infected, it is likely that all plants in the monoculture would be equally susceptible. As expected, the mass infection of monoculture crops has occurred numerous times. In the relationship between banana (Musa spp.) cultivars and the fungal pathogen *Fusarium oxysporum* f.sp. *cubense* (*Foc*) (the causal agent of Panama Disease/Fusarial vascular wilt), multiple lineages (or races) have reemerged (Randy C. Ploetz, 1994). Plantations of the easily internationally transportable banana cultivar Gros michel, which was susceptible to *Foc* race 1, were decimated in the 1960s. As a response, a new cultivar of banana called Cavendish was bred (Ordonez et al., 2015). The estimated financial cost associated with the massive loss of plantations from the 1960s totals around $2.3 billion (after inflation adjustment to the year 2000) (R. C. Ploetz, 2005). By the end of the decade, plantations of Cavendish in Taiwan had similar symptoms. Later, in 1994 these disease symptoms were attributed to *Foc* tropical race 4 (Ordonez et al., 2015). The Cavendish banana is yet to be replaced by a similar and easily transportable cultivar. Meanwhile, *Foc* tropical race 4 continues to spread into new territory such as Vietnam (Hung et al., 2018).
In another example of pathogen pressure on a monoculture system, tomato cultivars have also become susceptible to *Fusarium oxysporum* f. sp. *lycopersici* (*Fol* / *Fo4287*). Due to differences in breeding (compared to banana), single resistance genes (*I*, *I*-2, and *I*-3) were bred into cultivars to confer resistance to *Fol* races. The gap of time between the introduction and development of susceptibility of each new resistant variety was about 20 years in each case. (Takken & Rep, 2010). While the exact mechanisms of the expressed I proteins are mostly unknown, research into the specific domains of these proteins has revealed some answers, such as the activation states of I-2 (Tameling, 2006). Within the last ten years, a fuller picture of the molecular interaction between *Fol* and tomato resistance has been obtained (Takken & Rep, 2010). The actual interactions between pathogens and their host(s) are rarely as simple as the Arm’s Race cycle of *R* and *AVR* genes. Usually for progress to be made in understanding these interactions, examination is required from multiple angles. As a result of the loss of efficiency of many monoculture systems, different farming practices such as intercropping (Han et al., 2016) and crop rotation (Marburger et al., 2015) have been recently gaining popularity to assist in the control of disease causing pathogens. Due to the presence of pathogens with a wide host range such as *F. oxysporum* (with more than 100+ spp.) (Leslie & Summerell, 2006), different agricultural tactics do not always reduce the prevalence of the pathogen (Marburger et al., 2015).

How do new cultivars become susceptible to new races of *F. oxysporum* so quickly? The reason is twofold: first, the generation time of microbes is shorter than that of plants, and microbes have a high reproduction rate (Croll & McDonald, 2012). Second, in a tightly controlled system such as monoculture, genetically identical plants...
are used each growing period and therefore the host genome remains stagnant and unable to evolve resistance against pathogens that may be evolving the ability to infect.

In the interest of comparing pathogen-host interactions and to better understand the virulence of pathogens, the *F. oxysporum* species complex (*FOSC*), with its broad host range (Leslie & Summerell, 2006), can be employed as a model system for comparative analysis. Each isolate has a specific interaction with its host(s) which allows for comparisons to be made between pathogen and non-pathogenic *forme specialis*. Similarly to other filamentous plant pathogens (Dong, Raffaele, & Kamoun, 2015), *F. oxysporum* f. sp. *lycopersici*’s (*Fo4287*) genome has a high repeat content of 28% (Ma et al., 2010). The repeat-dense regions of plant pathogens are usually associated with higher evolution rates and virulence genes (Ma et al., 2010; Raffaele & Kamoun, 2012). Occasionally, transposable elements (TEs), the causal agents of the spread of repetitive sequence, are directly linked to effectors. With *F. oxysporum*’s secreted in xylem (SIX) effector genes, the TEs known as miniature impala elements (MIMP(s) are often located in the promoter region (Dam & Rep, 2017). In addition, there are lineage-specific (LS) regions associated with host specificity and related effectors in *F. oxysporum*. Contained in these LS regions are a majority of the TEs of the genome (Ma et al., 2010). To examine the implications that TEs may have for virulence, specifically with *Fusarium oxysporum*, a deeper look into TE function is required.

**Introduction**

From her research conducted on maize in the 1940s, Barbara McClintock discovered translocations of specific genetic elements (McClintock, 1950). McClintock believed these mobile controlling elements were the cause of varied color patterns in
maize kernels due to the ability to move between chromosomes. Though she was later awarded the 1983 Nobel Prize in Physiology or Medicine for her research on maize, the idea of genetic elements moving and taking part in regulation/development was still not widely accepted at the time (Biémont & Vieira, 2006). Today, we know these mobile genetic elements as transposable elements (TEs). In the past, they were most commonly described as “selfish and parasitic” and/or junk DNA. In the present day, TEs are viewed as another layer of complexity to the evolution and regulation of genomes.

**Classification**

TEs are composed of sequence(s) that allow them to multiply and move in the host genome, which allows them to be retained in the genome. The mechanism by which the TE sequence moves or transposes determines its highest level of classification. According to the classification system by Wicker *et al.* (2007), categorization of TEs into class I or class II is decided based on the presence or absence of an RNA intermediate respectively. In the simplest terms, class I TEs have a copy-and-paste-like mechanism and are called retrotransposons while class II TEs have a cut-and-paste-like mechanism and are called DNA transposons. The subdivisions within each of these classes are determined by a range of criteria such as: the specific sequences at the flanking ends of the TEs, components that are coded within their sequences (or lack thereof), the sizes of their target site duplications (TSDs), and more.

Within TEs there are the DNA sequences for proteins that give the TE the ability to replicate/excise and move to another region in the genome. TEs can be autonomous or non-autonomous depending on the components that are encoded in the TE sequence. As a general idea, a “complete” and active retrotransposon with autonomous function will
have a repertoire of proteins encoded in its sequence. These proteins include: a reverse transcriptase enzyme (RT), a ribonuclease H (RH), an integrase enzyme (INT), and a protease (PR). Through the transcription and translation of the TE sequence, the corresponding proteins are synthesized in the host cell. Included in this is a PR which cleaves the proteins and assembles them for the upcoming steps. From the TE sequence, the single-stranded RNA intermediate is transcribed. This RNA is then met with RT which can read RNA and create a double-stranded DNA (with the help of RH to degrade the RNA from the DNA-RNA hybrid that is made in the transposition process). The double stranded DNA is then able, through the action of INT, to be spliced into a new location in the genome where the TE will reside (Biémont & Vieira, 2006).

Class I elements are further organized into five orders: long terminal repeat (LTR) retrotransposons, short interspersed nuclear elements (SINEs), long interspersed nuclear elements (LINEs), Penelope-like elements (PLEs), and Dictyostelium intermediate repeat sequence (DIRS). Retrotransposons are classified into these orders based on the sequence configuration of their required components and variations in the mechanism that they use to replicate. Within each retrotransposon order are superfamilies which have a common/similar approach to their transposition. LTR retrotransposons are highly represented in plants. In non-Animalia organisms, Gypsy and Copia are the two primary superfamilies of the LTR order. The difference between these two superfamilies is the arrangement of their reverse transcriptase and integrase enzyme genes. Long interspersed nuclear elements, as their name suggests, can be relatively long - up to several kilobases. LINEs do not contain LTRs and if autonomous (having the ability to use their own components for transposition), the sequence of the LINE will at least contain a reverse
transcriptase and a nuclease. The five major superfamilies are categorized by their variation or presence of different components such as a endonuclease or RNaseH (Wicker et al., 2007).

Class II elements (unlike Class I elements) are initially divided into two subclasses. Subclass I is composed of order 1, which contains DNA transposons that move through the aforementioned “cut-and-paste” mechanism involving the cleavage of both strands of the DNA sequence. The division of subfamilies within this subclass is determined based on the sequence of their terminal inverted repeats (TIRs) and the length of their TSDs. Within the second order of subclass I reside the Crypton TEs. It is believed that they transpose without an RNA intermediate (due to their lack of RT) and instead use a tyrosine recombinase (YR) (Wicker et al., 2007). Due to their lack of TIRs, but cleavage of the double strands of DNA, these transposons retain their position in subclass 1.

Subclass 2 of the DNA transposons is composed of TEs that transpose without the cleavage of both DNA strands. Their transposition mechanism differs, and this subclass includes the Helitron order (which utilizes a “rolling-circle” mechanism) and the Maverick order (Feschotte & Pritham, 2007; Wicker et al., 2007).

Functions of TEs in Their Host

In 1980, while scientific acceptance of TEs was slowly gaining ground, the concept that they were only “selfish” DNA was contrived. In opposition to the fact that TEs have been retained in genomes over evolutionary time, it was argued that TE sequences do not result in a selective advantage for the organism in which they reside. The claim was that TEs survived through evolution only due to the transposition
mechanisms in their code which allowed them to persist. This extra DNA was therefore seen as an “energetic burden” (Doolittle & Sapienza, 1980). As research on TEs continued, however, this concept that TEs are “wasteful” was discarded. Due to the distribution of superfamilies in multiple kingdoms in the eukaryotic tree of life, it has become increasingly clear that TEs have ancient origins (and importance) (Feschotte & Pritham, 2007). It has been suggested these elements are responsible for “rewiring” existing regulatory systems are and a source of regulatory elements. This rewiring can occur when the same retrotransposon copies itself and “paste” to multiple locations (in the genome) in the regulatory flanking regions of genes. This TE can then perform a role as a regulatory motif for the affected genes to all be brought under the same regulatory system (Feschotte, 2008).

While TEs can have destructive results depending on where they insert, the change will only negatively affect the species if the changes occur in the germline. If the negative change reduces fitness in the offspring, then the organism will be less likely to pass on this variation of genetic material to the next generation. TE insertions can cause mutations which can cause disease but mutations are also part of creating variation in species (Biémont & Vieira, 2006). Acceptance of and interest in TEs increased after the sequencing of the human genome was complete in 2001 (Lander et al., 2001). The draft of the human genome resulted in the discovery that 45% of the DNA in the human genome is composed of TEs. In addition, this meant that many human genes were likely derived from TEs. With the evidence that TEs are not always associated with detrimental effect, many scientists to let go of the notion that TEs are selfish (Biémont, 2010; Lander et al., 2001). With this new information, the idea that suggested (before the draft of the
human genome) that TEs may have become “domesticated” in their host genome for a specific function became more readily acceptable as an explanation as to why TEs were so prevalent in some genomes (McDonald, 1983).

In the plant-pathogen interaction, a TE insertion could improve virulence or lead to the death of a failed variant. The consequence of a TE induced change in a pathogen’s sexual form could lead to the deactivation of an effector required for virulence over the host. This TE-induced avirulence would likely not be found in nature due to failed infection and spread. In the opposite way, virulence has been regained by pathogenic strains due to TE insertions. In the case of Magnaporthe oryzae, the fungal pathogen responsible for rice blast, transposon insertion into AVR-Pita, prevented the host from recognizing the previously avirulent protein, thus allowing for resistance (Zhou, Jia, Singh, Correll, & Lee, 2007).

Although the results could be detrimental, depending on the position of the insertion it has been found that most insertions on their own are neutral (Biémont & Vieira, 2006). While the cause of the differences is not well understood the rate of DNA mutation induced by TE insertions differs between organisms. Some of these insertions are a major source of new genetic material for the organisms affected. This rate ranges from 50-80% (among individuals) in Drosophila melanogaster and from 0.1%-1% in the human genome (Biémont & Vieira, 2006). An alteration of a gene’s function (or regulation) due to a TE insertion could lead to the (albeit rare) development of disease in humans. It has been found that such alterations are the cause of 0.5-1% of human illnesses (Kazazian, 1998). Health problems associated with the insertion of SINEs and LINEs near specific genes include: breast cancer, tumors of the esophagus and
reproductive organs, hemophilia, and Duchenne muscular dystrophy (Biémont & Vieira, 2006).

**Exaptation of TEs and Genome Organization**

TEs can create additional networks of regulation; insertions can, for example, initiate DNA methylation or histone modifications. These modifications can cause the DNA regions that flank the TE to become silenced in addition to the TE also becoming repressed. A TE insertion which initiates the formation of heterochromatin could be co-opted as an epigenetic marker dependent on environmental conditions (Feschotte, 2008). It was previously believed that epigenetic modification of TE sites came about as a defense against TE-introduced changes, to prevent the potential negatives brought about by a TE insertion. If this defense hypothesis was correct, then particularly younger and more likely to be active TEs would primarily have marks for repression. Chromatin immunoprecipitation mapping of 38 histone modifications to their associated sequence tags in human CD4+ T cells showed that there are both repressive and active histone modifications targeted to TEs and that older TE families were usually more associated with histone modifications. These results support the exaptation hypothesis - that specific TEs have been exapted for a function in the host genome in which they reside and through evolution have retained these changes (Huda, Mariño-Ramírez, & Jordan, 2010).

While this exaptation model may be supported by the evidence with human CD4+ T cells, the functions of TEs may not be identical in other organismal systems. As suggested by Seidl & Thomma (2017), in many fungal genomes, regions with high levels of TEs are usually epigenetically regulated due to markers for heterochromatin. Specifically, *Fusarium graminearum*, most commonly known for its ability to cause
disease on many agriculturally important cereal grains such as wheat and maize, has regulated gene clusters associated with its pathogenicity. These secondary metabolite (SM) gene clusters are also in TE-concentrated regions a location which results in chromatin-regulated expression (Connolly, Smith, & Freitag, 2013; Seidl & Thomma, 2017). The reorganization of the genome by TE-directed translocation along with rewiring of regulatory networks (Feschotte, 2008) allows some fungus to gain a more efficiently organized genome. If genes in a SM pathway are clustered, then survival of the pathway into another species (or same species) of fungi after a horizontal transfer event is more likely. If the genes of an important (pathogenicity-related) pathway were separated, then horizontal transfers would likely result in only (nonfunctional) fragments of the pathway (Walton, 2000). In addition to the unique clustering of SM genes in fungi, the chromosomal location seems to also be important. It has been shown that clustering often occurs at sub-telomeric regions (Palmer & Keller, 2010).

The location of pathogenicity related genes is also very important in various *Fusarium* species. In a genomic comparison between *Fusarium oxysporum* f. sp. *lycopersici* (*Fo*4287) (a tomato pathogen strain) and *Fusarium verticillioides* (a maize pathogen strain), more than 90% of the genome has synteny. These syntenic regions are composed of “house-keeping genes” that are required for *Fusarium’s* general upkeep such as growth and metabolism. *Fusarium oxysporum* also contains lineage-specific (LS) chromosomal regions which are related to its ability to infect its host(s). In the case of *Fo*4287, the LS regions are unique sequences which are composed of 40% of the genome (Ma et al., 2010). The LS regions of *Fo*4287 are generally sub-telomeric or at the ends of
chromosomes, and unsurprisingly 74% of the TEs in the genome are in this region (Ma et al., 2010). As was suggested by Walton (2000), the clustering of SM genes increases the likelihood of the full horizontal transfer of the cluster between species. In *Fo*4287, it was hypothesized by Ma et al. (2010) that horizontal transfer could also be an origin of the LS regions. Due to the association between TEs and regions related to virulence/pathogenicity, further research is required in fungal evolutionary biology to find the direct effect, if any, that TEs have on the infection course and pathogenicity.

The *F. oxysporum* species complex has a large host range (Leslie & Summerell, 2006), with isolates that have a very specific host - pathogen interaction. Due to the LS regions having a high density of TEs (Ma et al., 2010), *FOSC* can be utilized as a model system to find the relationship between TEs and virulence and the infection course. In this study, both RNA sequence data (for expression data) and genomic data were analyzed to draw correlations concerning TEs and the infection of *Arabidopsis thaliana* by two *F. oxysporum* isolates.

**Methods**

To examine the transposable element content of two *Fusarium oxysporum* genomes, those of *Fo*5176 (*A. thaliana*-infecting isolate) and *Fo*47 (non-pathogenic isolate), DNA sequencing reads were extracted from the National Center for Biotechnology Information (NCBI) sequence read archive (SRA). Genomic sequences for *Fo*47 came from all runs of multiple submission samples (Accessions: SRX101559, SRX101565, SRX101570, SRX081457, SRX081458, SRX081479, SRX081482, SRX081484) and were all from Illumina HiSeq 2000 paired-end sequencing. Genomic sequences for *Fo*5176 were represented by single end reads (Accessions: SRR305330,
SRR305356, SRR305359, and SRR305380). File extraction was performed using SRA toolkit (version 2.8.0) (“Sequence Read Archive Toolkit,” 2011) and files were dumped into the Massachusetts Green High Performance Computing Center (MGHPCC) cluster server for workflow including file conversion and mapping. A TE library sequence fasta file was made for each strain composed of TEs, effector (SIX) genes, and ten control genes (different for each strain). The ten control genes were selected based on their consistent expression (lowest variance from a previous analysis by Li Guo, unpublished data) throughout all time points. For future analysis, copy number data on the *Fusarium* strain *Fo*4287 (tomato pathogen strain) was also generated (unpublished data).

For each strain, the following modules and methods were used. After file extraction, the bowtie2 module (version 2.3.2) (Langmead, et al., 2011) mapped the concated (fastq) whole genome sequence reads to the TE library sequences file. Next, the SAMtools module (version 1.4.1) (Li, et al., 2009) was used for file conversion (from sam to bam) and to filter out reads that did not map to the TE library. The bam file was sorted and indexed. The sorted files were then used by the bedtools module (version 2.26.0) (Quinlan & Hall, 2010) to calculate the coverage (with nucleotide resolution). To normalize the copy number of the sequences from the concatenated fastq file(s), the median coverage had to be calculated. From each strain, the coverage of ten control genes that each had a single copy was normalized by the corresponding gene length. The values for the 10 control genes were then averaged. The raw coverage for every TE in each strain was then normalized by dividing by the average value.

On MatLab (release R2017a) (“MATLAB,” 2017), clustering was done (with a k value of 7). This clustering grouped the TEs based on their copy number differences
Clean-up and Validation

Modifications were made to the custom script to remove outliers due to mapping error (fragments) when there were cases of copy numbers below 0.1. The custom script also removed the TE out of the library when none of the three tested strains (Fo47, Fo5176, and Fo47) had any copies of it. There were originally 140 TEs which included retrotransposons (such as SINEs, LINEs, and LTRs) and DNA transposons (such as hATs, pogos, Helitrons, Mariners, Mutators, and cryptons). Most notably, LINEs were removed through filtering processes (either through lack of copies in all species or copy numbers below 0.1). To estimate the copy number for SIX genes, a local BLAST database was made using extracted gene fasta files from the Broad Institute (Cambridge, Massachusetts) for Fo47 and Fo5176 (http://archive.broadinstitute.org/ftp/pub/annotation/fungi/fusarium(genomes/)). For some SIX genes, short fragments were mapped but no other evidence indicated the existence of the complete gene in the particular strain. A local BLAST search was done against all SIX genes to check if copy number values less than one were false positives. This allowed for another check to be made. SIX genes were also utilized for validation after normalization. Since the copy number is already known, any copy number values that did not match led to further debugging of the custom code.

To examine the expression levels of TEs within Fo47 and Fo5176 during the infection of Arabidopsis thaliana, an RNA sequencing data set (unpublished data by Li Gou) was used. RNA extraction/sequencing was done in triplicate at five time points: at 0
hours (before infection occurred), 12 hours post infection (hpi), 24 hpi, 48 hpi, and 96 hpi with each of the two strains. Sequencing resulted in paired-end reads. For each of the time points, there were at most three replicates. Using the MGHPCC cluster and software mentioned above, each sample’s (strain, hpi, replicate) paired-end files went through a similar process as described for the genomic data. The paired-end fastq files for each sample were mapped to the same TE library as with the genomic data with bowtie 2. The output sam files were then sorted and converted to bam files using samtools. The output alignment files were then converted to raw coverage files to be used in MatLab. In MatLab, the expression was normalized by using the same ten control genes with consistent expression in each of the two strains, to determine the geometric mean for every TE for every replicate at every time point. The geometric mean value was averaged across the three replicates. After normalization of expression, MatLab allowed for easy filtering and removal of any TEs that had no expression for both strains. In addition, a t-test was used to find the time points with an insignificant expression change (compared to the 0 hpi value). To retain their insignificance, these expression values were changed to their 0 hpi expression value. Values were then plotted onto a custom heatmap. Since expression is dependent upon the number of copies in the species, we utilized the clustering order from the generated copy number data (of Fo47, Fo5176, and Fo4287) and applied it to the ordering for expression.

**Results & Discussion**

**Clustering Based on Copy Number**

The copy number data of the three strains (Fo47, the non-pathogenic isolate Fo5176, the *A. thaliana*-infecting isolate, and Fo4287 the tomato pathogen) were
clustered into five groups and plotted in figure 1 based on the similarity and range within each cluster.

**Figure 1:** Plot of all 50 TEs after computational workflow. Copy number resides on the top half while the expression heatmap of Fo5176 and Fo47 lies on the bottom half. Clusters are separated by magenta vertical lines and based on pattern similarity. The copy number plot’s vertical axis is in logarithmic scale. The expression heatmap’s vertical axis is in hours post inoculation with the range of the heatmap going from 0 to 15 (normalized expression values).

### Clustering analysis

Group one contains 35 TEs which have the lowest copy number interval (0.1:32). For the most part, the TEs in this group have a common pattern where Fo47 has the lowest copy number and Fo5176 has the highest copy number (20 out 35 times compared to Fo4287’s 14/35 times).

Cluster two has one TE: HopAy267761. Out of the all the TEs in the TE library, it has the highest copy number at 189 copies (for Fo5176). Fo47 does not have this TE and Fo4287 has 3 copies of it. This TE has the most dramatic difference in copy number between the two pathogenic strains.
Cluster three has three TEs in it. These TEs are Hornet-small, FoxyAJ250814, and Foxy2. The first two TEs are shared only by the two pathogenic strains. In this group, Foxy2 is the only TE that is shared by all three strains. The copy number difference (for each of the three TEs) and average copy number difference (99 copies) between Fo5176 and Fo4287 are the highest out of all the multi-TE clusters. For Fo4287, the cluster’s range is ~64 copies of Hornet-small to ~127 copies of the SINE, FoxyAJ250814.

The fourth cluster contains 11 TEs. This is the second-largest cluster and is different from cluster one in that all copy numbers are shifted up. The minimum number of copies for a TE (between the three strains) in this cluster is ~2 copies (Fot1-M24 of Fo4287). In this cluster Fo5176 has the highest number of copies in all 11 TEs. Out of the 11 TEs, Fo47 does not possess four: NhORF4-like, Fo1, Fot1-M24, and Han-full. Whenever Fo47 does contain the respective TE, this non-pathogen strain always has the lowest copy number in this cluster. For the TEs in this group, Fo5176 has copy numbers almost always above the copy numbers of cluster one.

Cluster five is similar to cluster two in that it is the other single-TE cluster. The TE in this cluster is Fo-Helitron2-FOXG14222, which is present in far more copies (in Fo5176) than the other Helitrons clustered in group one. This TE has one of the largest copy number differences (~95 copies) though not nearly as large as the difference in cluster two (~185 copies). The reason this TE is not included in cluster three (with a similar copy number difference) is that Fo5176 dominates Fo4287 in this cluster, whereas cluster three shows the inverse pattern.
Correlation Between Copy Number and Expression

To find whether a correlation was present between the expression and copy number of TEs in \textit{Fo47} and \textit{Fo5176}, plots were made with copy number on the x axis and expression for 0 hpi on the y-axis (Figure 2 for \textit{Fo5176} and Figure 3 \textit{Fo47}).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{Plot of \textit{Fo5176} TE expression at 0 hpi and genomic copy number.}
\end{figure}

\textit{Fo5176}
Figure 3: Plot of Fo47 TE expression at 0 hpi and genomic copy number.

The correlation between the expression (0 hpi) and copy number is weak for both Fo47 (Figure 3) and Fo5176 (Figure 2). The Pearson correlation coefficient values are 0.274669 and 0.091468 for Fo47 and Fo5176, respectively. This shows that at least with the TEs from the library that were tested, there is a weak, positive correlation between the number of copies and the level of expression.

To find whether the expression level of the tested TEs correlates with the infection course of A. thaliana, a plot was made for each TE with expression on the y axis and time (hpi) on the x-axis. Error bars (with standard deviation) are used to draw significance from changes in expression. One plot was made (for each strain with expression data) to contain all 50 TEs examined (Figure 14 for Fo5176 and Figure 15 for Fo47). For each strain, the overall expression profiles of all 50 TEs were inconsistent,
though there were patterns that groups of TEs followed. The TEs that followed specific
patterns are listed in Table 1 for *Fo5176* and Table 2 for *Fo47*. TEs that are not present in
the table had no distinguishable pattern from their expression level at 0 hpi.

**Table 1:** Contains the observed patterns repressed expression of all tested TEs for
*Fo5176*. For each pattern, the TE name in bold is the TE shown in the respective plot.
The TE expression level that does not involve the interaction with the host is shown with
a “c” for control. Plots have an x-axis in hours (post inoculation) and y-axis normalized
expression. All expression plots are in Figure 14. Error bars indicate the standard
deviation of the samples.

<table>
<thead>
<tr>
<th>Pattern</th>
<th>TEs included</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Helitron-1, Helitron-2, Helitron-4, Frodo, Strider, Fot1, Fot2, Fotnew, Fot1-M24, Fot2-M24" /></td>
<td>Helitron-1, <strong>Helitron-2</strong>, Helitron-4, Frodo, Strider, Fot1, Fot2, Fotnew, Fot1-M24, Fot2-M24</td>
</tr>
<tr>
<td><img src="image" alt="Helitron-3, DrifterAY, Fot3, Fot3-M24, NhORF4-like, HopAY" /></td>
<td><strong>Helitron-3</strong>, DrifterAY, Fot3, Fot3-M24, NhORF4-like, HopAY</td>
</tr>
<tr>
<td><img src="image" alt="FoxyAJ, Foxy2" /></td>
<td><strong>FoxyAJ</strong>, Foxy2</td>
</tr>
<tr>
<td><img src="image" alt="Hornet" /></td>
<td>Hornet</td>
</tr>
</tbody>
</table>
Table 2: Contains the observed patterns induced expression of all tested TEs for Fo5176. For each pattern, the TE name in bold is the TE shown in the respective plot. The TE expression level that does not involve the interaction with the host is shown with a “c” for control. Plots have an x-axis in hours (post inoculation) and y-axis normalized expression. All expression plots are in Figure 14. Error bars indicate the standard deviation of the samples.

<table>
<thead>
<tr>
<th>TE Name</th>
<th>Expression Pattern</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helitron-5, Tfo1AB, Tfo1, Tfo2, Sam, Han-full</td>
<td><img src="image1.png" alt="Plot" /></td>
<td>Shows a sharp decrease during the infection course.</td>
</tr>
<tr>
<td>MarsuAF</td>
<td><img src="image2.png" alt="Plot" /></td>
<td>Shows a decrease in expression compared to the control expression level.</td>
</tr>
<tr>
<td>Tfo3, Gollum-full</td>
<td><img src="image3.png" alt="Plot" /></td>
<td>Shows a repression of expression pattern of Hornet, a DNA transposon, expression</td>
</tr>
</tbody>
</table>

There were two types of expression patterns for TEs in Fo5176, most of which involved a decrease in expression compared to the control expression level (Table 1). These patterns of suppression during infection differed in their period of increase, expression peak, and the rate that they decreased. The pattern related to the two short interspersed elements – FoxyAJ and Foxy2 was a sharp decrease during the infection course. For the repression of expression pattern of Hornet, a DNA transposon, expression
was fully repressed across all time points. For the induced expression patterns, shown in Table 2, there are less TEs that show an increase in expression compared to the control expression values. The TEs that do have induced expression, go back to (about) the same amount of expression of the control by the last time point (96 hpi). Tfo3, the DNA transposon and Gollum-full, the long terminal repeat (LTR – retrotransposon) are the only TEs that have an expression higher than the expression of the control across all time points.

Fo47 had fewer patterns (Figure 15) and an even lower total number of TEs were involved in any pattern or had any significant change in expression over the course of infection. Some TEs were grouped into the same pattern in both strains such as the pattern by FoxyAJ and Foxy 2 and the pattern shown by Fot2 and Fot2-M24. The only notable expression was that of FoxyAJ, Foxy2, and Helitron-2 all with a peak in expression at 48 hpi. As seen in the overview of copy number and expression (Figure 1), Fo47 has lower expression throughout all TEs. It is unclear if the lower expression in Fo47 was in part due to its low copy number.

Discussion:

As shown by previous research, transposable elements hold significant weight in evolutionary biology. The ability for TEs to rewire existing regulatory systems and unify them under common motifs can allow organisms with a high reproductive rate to quickly adapt to new environments (Feschotte, 2008; Huda et al., 2010). Due to the relevance of TEs in the genomes of pathogens (Dong et al., 2015), further bioinformatics analysis to better understand increased mutation rates, adaptation, and the association with TEs is needed. Finding the specific TE motifs associated with the LS chromosomes of F.
oxysporum could lead to further discoveries related to the horizontal transfer of genomic sequences between organisms.

This analysis showed that there was no single expression pattern of TEs. Nearly all TEs failed to show induced expression in the presence of a plant host. In Fo5176, only Tfo3 and Gollum-full had an increase in expression which remained above control levels at 96 hpi. MarsuAF, the only other TE that had increased expression in the interaction with the host, decreased to the control expression level by 96 hpi. In the case of Fo47, the two expressed SINEs: FoxyAJ and Foxy2 had expression levels higher than basal by 96 hpi, though the significance of this increase is minimal. No other TE in Fo47 had a notable increase in expression.

For the TEs that did express differently than the expression in the absence of the plant host, the function of this expression is not known. An experimental focus on TEs that are highly expressed in Fo5176, such as the SINE FoxyAJ, the DNA transposon mutator HopAY, and Helitron2, is needed before a function can be suggested. Though it is still unknown what the exact effects are of TEs in the infection course between A. thaliana and F. oxysporum 5176 and 47, distinct patterns of TE expression occur, more so in the pathogenic isolate. As two additional layers of computation, identifying the location of TEs of interest and the ratio of TE copies that are active could further elucidate the role of TEs in respect to virulence/pathogenicity. By finding the location of specific TEs, cis-regulation can be inferred if a TE is within the same region and has a similar expression pattern to known SM cluster genes, effectors, and other virulence factors. By finding the active ratio of a specific TE, the copy number can be better represented and therefore the expression per copy can be confidently reported. Due to the
wide host range of isolates within *FOSC*, there are many more comparisons that could be done with a similar analytical pipeline and an expanded TE library.
CHAPTER 2
PEPPER EXTRACTS AND GROWTH INHIBITION
OF **F. OXYSPORUM**

**Introduction**

Capsaicinoids are a group of alkaloids found in a majority of peppers within the genus, *Capsicum*. It has been suggested that capsaicinoids, and primarily capsaicin (8-Methyl-N-vanillyl-trans-6-nonenamide) have been utilized by pepper plants as a defense mechanism against fungi (Srinivasan, 2016) and mammals (Tewksbury & Nabhan, 2001). In nature, consumption of capsaicin-containing peppers by mammals results in the activation of vanilloid receptor 1 (VR1), a pain-sensing receptor (Julius et al., 1997). The resulting burning sensation has the ability to deter seed predators such as mammals, which do not provide any evolutionary advantage to peppers, because of seed predation (destruction of seeds) or weak dispersal (Schupp, Jordano, & Gómez, 2010). As opposed to mammals, birds are insensitive to capsaicinoids, though birds do have an ortholog of mammalian VR1 (Jordt & Julius, 2002) which makes them comparably suitable seed dispersal agents (Tewksbury & Nabhan, 2001). Due to the antimicrobial properties of capsaicin and its association with protecting seeds, it was used in ancient Mayan medicine for a range of ailments (Cichewicz & Thorpe, 1996). This further supports research showing capsaicin to have antimicrobial activity against Group A Streptococci (Marini, et al., 2015) and growth inhibition against *Fusarium semitectum* (Tewksbury et al., 2008).

After growing capsaicin containing chili pepper cultivars of *Capsicum chinense* from seed, extractions with 90% ethanol were done on whole fruits to retrieve a capsaicin-containing fraction. The fraction was confirmed to contain this alkaloid with
the use of high performance liquid chromatography (HPLC). These extracts were then used in screens against four strains of \textit{Fusarium oxysporum} using a 96 well plate and optical density reading for growth inhibition quantification.

**Methods**

**Initial Screens and Set-up**

To measure the inhibition of \textit{F. oxysporum} growth, a plate reader was used to measure optical density (OD) at 600 nm. For initial screens (while pepper plants were growing), \textit{F. oxysporum} strains \textit{Fo47} (a non-pathogenic isolate), \textit{Fo5176} (an \textit{A. thaliana} infecting isolate), \textit{Fo32931} (an isolate pathogenic to immunocompromised humans), and \textit{Fo4287} (an isolate pathogenic to tomato) were grown for spores to later be filtered. In general, all strains were grown in potato dextrose broth (PDB) (infusion of potatos and dextrose – 24 g/L deionized water) and shaken at 130-150 rpm at 27°C for about five to seven days to yield enough spores. Spores were filtered with several layers of miracloth and were counted with the use of a hemocytometer. Dilutions suitable for the screens were done based on the stock spore concentration.

In each screen, a clear-bottom 96-well plate was utilized. All wells had a total volume of 200 uL. Each column was a control or treatment. Controls included: untreated growth, ethanol (EtOH), and no \textit{F. oxysporum}. Each (non-control) well was given 176 µL of PDB, 20 µL of \textit{F. oxysporum} culture, and 4 µL of treatment. The untreated growth control was always the first column of the plate and was composed of 180 µL of PDB and 20 µL of culture. The ethanol control was composed of 176 µL of PDB, 20 µL of culture, and 4 µL of 100% ethanol. The control wells without \textit{F. oxysporum} (the bottom 2-3 rows of every column) contained 196 µL of PDB and 4 µL of the treatment
corresponding with the column’s treatment. These wells should show little to no change in their reading due to having no living organisms inside. In initial screens, the stock culture was diluted to a spore concentration of $1 \times 10^4$ to $1 \times 10^7$ spores/mL in PDB. The dilution of 20 µL of culture into the final volume of 200 µL brought the final concentration (range) of spores in screens to $1 \times 10^3$ to $1 \times 10^6$ spores/mL PDB. In screens using pepper extracts, a consistent (final) spore concentration of $\sim 1 \times 10^6$ spores/mL PDB was used. In initial screens, dilutions of capsaicin with a maximum concentration of 1 mg/mL were used for the 4 µL of treatment.

The plate reader was set to a temperature of 27 °C, and the absorbance of the entire plate was read every 30 minutes for 24 hours post inoculation. These 24-hour screens utilized the SpectraMax M5 by Molecular Devices (San Jose, CA). A screen was later done with a plate cover and at a temperature of 25 °C for 48 hours. This screen, too, used a SpectraMax M2 by Molecular Devices. Before each read, the plate was also set to shake for 3-5 seconds.

**Growing Chili Peppers**

Three different *Capsicum chinense* cultivars were grown: Habanero White Chili (WC), Red Moruga Scorpion Chili (MSC), and Carolina Reaper Chili (CR). Seeds came directly from whole peppers that were gifted by Tracy James Powell. They were cut out of the peppers and surface-sterilized before being germinated in an incubator. Carolina Reaper Chili seeds were also bought from *Seeds & More* (Newfoundland and Labrador, Canada) leading to four pepper plant groups. To avoid confusion, the Carolina Reaper Chili pepper seeds from *Seeds & More* and their derivatives will be referred to as CRC, to distinguish these seeds from the CR seeds donated by T.J. Powell. Pepper seedlings
began in a Percival (AR41L2) growth chamber in the Life Sciences Laboratory with a 16-hour light cycle and temperature of 23-25 °C. These plants were watered every other day and fed 1 tablespoon of (15-30-15 NPK) nutrients diluted in one gallon of water every other week. The pepper plants remained in the growth chamber for around three months until they required more space. The largest eleven pepper plants of the original fourteen plants were transferred to the CNS greenhouse to continue growing. The greenhouse conditions were more closely regulated, with a 14-hour light cycle and a daytime temperature of 22 °C and a night-time temperature of 18 °C. During each watering, the plants received 200 ppm of nitrogen (Peters Professional 20-10-20 NPK peat lite special).

Upon flowering, each cultivar was self-pollinated with the use of paintbrushes. The harvest of peppers was forced earlier due to thrips and a potential virus. Samples were tested by (Angie Madeiras at) the University of Massachusetts Amherst Plant Diagnostics lab for impatiens necrotic spot virus and tomato spotted wilt virus, both of which came back negative. While there were about 30 flowers per plant upon self-pollination, four of the eleven plants did not produce fruit. Peppers were harvested and put into labeled bags. They were later rinsed, air dried, and refrigerated at 4°C until their extraction.

CRC produced few chilis in the initial harvest and was surpassed by CR and WC (data not shown). After the viral tests came back negative, a second harvest was done about 3 weeks after the initial harvest. This harvest resulted in many CRCs to around the number of CRs in the first harvest. The chilis from the second harvest of CR, WC, and MSC were not used in any extraction.
Extraction of Capsaicin

To retrieve a capsaicin-containing pepper extract from fresh peppers, peppers must be ground, and a solvent used to extract the capsaicin. Due to four different sample cultivars being used and the strenuous pestle-and-mortar grinding, pepper extractions occurred on different days. Following the extraction procedures of Chinn et al. (2011), 90% ethanol was used as the solvent at 15% w/v of the fresh peppers. After weighing, peppers were broken up in a fume hood (by hand) into smaller parts and put into a pestle and mortar with liquid nitrogen. The peppers were ground until they were a fine powder which was then added to a flask of the (premeasured) solvent. After all peppers were ground, the extraction flask was put on a shaker for 24+ hours, though according to previous analysis by Chinn et al. (2011), after 24 hours, levels of extracted capsaicin did not change dramatically. The following day, the flask was taken off the shaker and allowed to sit at an angle for at least 24 hours to allow the solids to settle. Using a glass Pasteur pipette, the top layer of the liquid in the extraction flask was taken and deposited into glass culture plates (and a single, 1-mL glass HPLC vial) for drying in a fume hood. After several days of drying, the plates were scraped with a razor blade, and the dried residue was put into HPLC vials and refrigerated (at 4 °C). Additional HPLC vials that had each received 1 mL of unfiltered liquid extract were refrigerated to be used with HPLC for later comparison once the ethanol had evaporated.

As a test, 600 mg of a dried extract (scraped from a glass culture plate) was added to 100% ethanol in which the extract did not fully dissolve. To avoid this problem, each of the scraped extracts was dissolved in an appropriate amount of 100% ethanol (100 mg extract/mL ethanol) and cleaned using vacuum filtration (with a Buchner flask and funnel) and paper filter disc. The extracts were given time to dry in glass culture plates.
With all samples to be used in HPLC, plastic pipettes were substituted with glass Pasteur pipettes. In multiple Eppendorf tubes (and HPLC vials for HPLC tests), extracts were weighed, and 1 mL of 100% ethanol was added. All tubes had 50 or 100 mg of extract. Even after shaking and a suitable amount of time (more than 24 hours for all extracts), a solid was still precipitating out of solution. To retrieve the fraction most likely to contain capsaicinoids, the extracts were centrifuged (at 4000 rpm), and the top (liquid) fraction was transferred to new containers for each of the extract groups. To avoid confusion, the naming of extracts will be: the pepper abbreviation followed by the mass of the solid pepper extract that was used in the final step. As an example, a 100 mg CRC solid extract was added to a tube with 1 mL of 100% ethanol and mixed. The removed liquid fraction would therefore be designated “CRC 100ex” and be used in screens. To make a “25ex” sample, ethanol (100%) was used to dilute the “50ex” sample two-fold. The unfiltered extracts are designated with a “u” before the cultivar abbreviation.

**High Performance Liquid Chromatography**

To find the relative concentration of capsaicin in the pepper extract samples, the Agilent Technologies (Santa Clara, CA) 1260 Infinity II High Performance Liquid Chromatography (HPLC) was used with a Zorbax Eclipse Plus C18 column (5µm 4.6x150mm). The set-up included an automatic sampler set to an injection volume of 50 µL. The quaternary pump was set to a flow rate of 1 mL/min with a mobile phase of 70% acetonitrile and 30% autoclaved distilled water. The UV/vis detector was set to a wavelength of 280nm (capsaicin’s absorbance value). A 12-step serial dilution was done with 95% capsaicin obtained from Sigma-Aldrich (St. Louis, MO). Capsaicin with a purity of 95% was used because it was previously purchased for optical density growth.
inhibition screens and did not show any issues in the chromatograms of the HPLC. The dilution started at 1mg/mL 95% capsaicin and was diluted by half each time with 100% ethanol. Each of the 12 vials containing the dilution of capsaicin were analyzed by HPLC for 15 minutes.

The associated program, Agilent ChemStation, was used for data processing. The chromatograms were auto-integrated to create the calibration curve based on the area of the capsaicin associated peak. A peak at a retention time of ~6.6 minutes continually appeared on all but the last chromatogram (the lowest dilution) shown in Figure 4. The area of the peak at ~6.6 minutes for each standard sample (calculated by auto-integration) and presented in Table 3 was paired with the known (actual) capsaicin concentration of each sample to find the correlation and make a calibration curve. The points were plotted using Microsoft Excel; a linear line of best fit was made (R² = 0.9983) and is shown in Figure 5.

Following the same procedure as for the standards, the unfiltered and 50ex pepper extract samples were run through the HPLC. All generated chromatograms from the unfiltered extracts (Figure 16) and 50ex extracts (Figure 6) were processed in an identical manner as for the capsaicin standard chromatograms. The area of the capsaicin-associated peak at ~6.6 minutes for each sample was calculated by the auto-integration mechanic. Exact peak areas and retention times of 50ex and unfiltered extracts are listed in Table 5. Using the line of best fit, the amount of capsaicin present in each of the tested extract samples was estimated (Table 4 and Figure 17).
Optical Density Screens with Extracts

The retrieved liquid fractions described in the previous sections were used in optical density (OD) reading screens using 96-well plates and *F. oxysporum* spores in an identical manner as with the 95% capsaicin standard.

Computational Analysis

To start the growth curve at zero for OD screens, an average was taken of the first three data points for each replicate of each treatment or control. This average was used to subtract across all time points for each respective replicate of each treatment or control. All analysis was done with the use of MatLab. To capture the data of all replicates (as separate lines) and compare the treatments, figures were made for each screen, with one graph for each of the treatment or control columns.
Results & Discussion

HPLC Results

Chromatogram of Capsaicin Dilution for Calibration

**Figure 4:** Overlay of all 11 chromatograms from the 95% capsaicin standard. Y-axis is measured in milli-absorption units (mAU) and x-axis is measured in minutes retention time. The resulting chromatogram are from a 1:2 serial dilution with 95% capsaicin and ethanol. Corrected capsaicin concentrations on Table 4. Capsaicin is responsible for peak at ~6.6 min retention time. Peak was present at ~6.6 min for all but lowest dilution. Areas from the 11 peaks at ~6.6 min were used for standard curve (Figure 5).

**Table 3:** This table refers to the chromatograms of Figure 4. These are the areas calculated by auto-integration for each of the peaks at ~6.6 minutes from the capsaicin standard serial dilution. Capsaicin concentrations were corrected for actual capsaicin in the 95% standard. Area is measured in milli-absorption units (mAU) multiplied by time. Exact retention times are also shown in the left-most column. Area values and capsaicin concentrations shown here are represented in Figure 5 in a calibration curve.

<table>
<thead>
<tr>
<th>Capsaicin (mg/mL)</th>
<th>Area (under peak)</th>
<th>Retention Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.95</td>
<td>24666</td>
<td>6.644</td>
</tr>
<tr>
<td>0.475</td>
<td>11248</td>
<td>6.589</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>0.2375</td>
<td>5769.2</td>
<td>6.593</td>
</tr>
<tr>
<td>0.11875</td>
<td>2967.5</td>
<td>6.597</td>
</tr>
<tr>
<td>0.059375</td>
<td>1534.7</td>
<td>6.602</td>
</tr>
<tr>
<td>0.0296875</td>
<td>780.73</td>
<td>6.596</td>
</tr>
<tr>
<td>0.01484375</td>
<td>404.55</td>
<td>6.602</td>
</tr>
<tr>
<td>0.007421875</td>
<td>202.01</td>
<td>6.6</td>
</tr>
<tr>
<td>0.003710938</td>
<td>108.22</td>
<td>6.609</td>
</tr>
<tr>
<td>0.001855469</td>
<td>79.424</td>
<td>6.644</td>
</tr>
<tr>
<td>0.000927734</td>
<td>18.422</td>
<td>6.605</td>
</tr>
<tr>
<td>0.000463867</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**Figure 5:** Calibration curve of capsaicin serial dilution. Exact area values are shown in table 4. Black line connects each of the serial dilution area values. Dotted red line is linear line of best fit. Maximum capsaicin standard was 1 mg/mL capsaicin (0.95 mg/mL actual). Equation for line of best fit is on right side of the figure with an $R^2$ value of 0.9983.

According to the HPLC chromatograms (Figure 6), out of all the 50ex samples, MSC has the highest level of capsaicin. WC has the lowest amount of capsaicin. The two Carolina Reaper Chili samples (CRC and CR) from different origins have very similar levels of capsaicin. The peaks at other retention times are unknown and would require the
use of other standards or other identification procedures such as gas chromatography-mass spectrometry. Due to limited amounts of extract, only one 50 mg sample of each cultivar extraction was placed into an HPLC vial for analysis. Carolina Reaper Chili was ranked as the hottest pepper at the start of this project, taking the title from Red Moruga Scorpion Chili. Outside of CRC, all peppers were harvested once and at the same time. It is expected that all cultivars are different and therefore have different ripening times, though further ripening does not always result in an increase in the levels of capsaicinoids (Nagy, Daood, Ambrózy, & Helyes, 2015; Nugroho, 2016). Outside of CRC, which was a defined lineage obtained from a seed-supply company, it is also unknown if these C. chinense cultivars (CR, MSC, and WC) are pure-breeding. The growing conditions of the parental plants that yielded the peppers/seeds (from Tracey Powell) are unknown. Crossing between other cultivars may have occurred, and the seeds used may not have been pure CR, MSC, or WC. Due to the observed low pollination rate (data not shown) of these chilis except WC, it is unlikely that they naturally cross-pollinated in the indoor environment of my project.
**Figure 6:** Y-axis is measured in milli-absorption units (mAU) and x-axis is measured in minutes retention time. Extract sample is listed on the left. WC = Habanero White Chili, CR= Carolina Reaper Chili, MSC = Moruga Scorpion Chili, CRC = Carolina Reaper Chili. These chromatograms are from the 50ex corresponding extract. Arrows designates capsaicin associated peak. As shown by chromatograms of standard capsaicin, (figure 4) capsaicin is responsible for peak at ~6.6 min retention time.
Table 4: This table refers to the chromatograms of 50ex (figure 6) and chromatograms of unfiltered extracts (Figure 16). Listed here are the areas calculated by auto-integration for each of the peaks at ~6.6 minutes. Area is measured in milli-absorption units (mAU) multiplied by time. Exact retention times are also shown in the third column. The estimated amounts of capsaicin for each of the samples are listed in the right most column in mg/mL. Each sample was in 1 mL of 100% ethanol.

<table>
<thead>
<tr>
<th>Extract Samples</th>
<th>Area</th>
<th>Retention Time</th>
<th>Estimated Capsaicin (mg/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WC 50ex</td>
<td>8495.2</td>
<td>6.589</td>
<td>0.33510957</td>
</tr>
<tr>
<td>CR 50ex</td>
<td>19704.8</td>
<td>6.604</td>
<td>0.77377279</td>
</tr>
<tr>
<td>MSC 50ex</td>
<td>113705</td>
<td>6.657</td>
<td>4.4522654</td>
</tr>
<tr>
<td>CRC 50ex</td>
<td>23806</td>
<td>6.607</td>
<td>0.9342643</td>
</tr>
<tr>
<td>uWC</td>
<td>429.8</td>
<td>6.597</td>
<td>0.01948775</td>
</tr>
<tr>
<td>uCR</td>
<td>1900</td>
<td>6.597</td>
<td>0.07702082</td>
</tr>
<tr>
<td>uMSC</td>
<td>22728.7</td>
<td>6.603</td>
<td>0.89210652</td>
</tr>
<tr>
<td>uCRC</td>
<td>11223.9</td>
<td>6.595</td>
<td>0.44189129</td>
</tr>
</tbody>
</table>

Optical Density Screens with Extracts

Through many screens it was shown that the most representative results are seen when the OD reading is done with the plate covered. If the plate was uncovered, plateauing of growth almost always occurred, regardless of the effect of the extract. Due to low final OD reads after 24 hours, a proper gap in inhibition could not always be seen (data not shown). The 48-hour screen of Fo32931 (Figure 18) and Fo4287 (Figure 19) had the clearest results for differences in growth. As a better overview of the effect of all extracts and to represent growth inhibition in another form, the total area under each of the growth curves was calculated and put into a bar graph for comparison between extracts and controls. (Figure 20 for Fo32931 and Figure 21 for Fo4287).

By using the area under the curve, each bar representing treatment was normalized to the amount of growth in the ethanol control. Each area associated with a treatment was normalized to ethanol instead of untreated growth because all extracts were
dissolved in 100% ethanol. This normalization was done by dividing the area under the curve of each of the extracts by area under the curve for the ethanol control. These respective fractions were then turned into percent inhibition after subtracting from 1 and multiplying by 100%. These values for Growth Inhibition (%) were put into bar graph form as well for Fo32931 (Figure 7) and Fo4287 (Figure 8). Fo32931 shows a higher sensitivity to the extracts compared to Fo4287 as shown with a higher percent growth inhibition with each of the extracts. There are also more inhibition values with significance in the Fo32931 data set. In the two growth inhibition data sets, Moruga Scorpion Chili extracts are the most growth inhibiting. Specifically, 50ex has ~75% inhibition against Fo32931 and about 50% inhibition against Fo4287.

![Figure 7](image_url)

**Figure 7:** Growth inhibition of Fo32931 for each of the pepper extracts, nystatin, and 95% capsaicin standard (x-axis). Areas under the curve for for each extract and control (Figure 20) were normalized by the area under the curve of growth in ethanol and converted to percentage form (y-axis). A t-test was done for each treatment in comparison to ethanol growth to find significance of growth inhibition. All significant values are shown with an asterisk.
The spore concentration of each well (with culture) was ~1.26E6 spores/mL. Plate was incubated at 25 °C and was covered.

**Figure 8:** Growth inhibition of *Fo4287* for each of the pepper extracts, nystatin, and 95% capsaicin standard (x-axis). Areas under the curve for each extract and control (Figure 21) were normalized by the area under the curve of growth in ethanol and converted to percentage form (y-axis). A t-test was done for each treatment in comparison to ethanol growth to find significance of growth inhibition. All significant values are shown with an asterisk (p<0.05). The spore concentration of each well (with culture) was ~1.25E6 spores/mL. Plate was incubated at 25 °C and was covered.

These two data sets are from a single 48-hour screen in which the 96-well plate was divided between the two isolates (*Fo32931* and *Fo4287*). In previous screens, 95% capsaicin at final (well) concentrations of 1 mg/mL and 0.5 mg/mL gave results that could not be interpreted (Figure 22). Due to the lack of solubility of capsaicin in water (“Hazardous Substances Data Bank: Capsaicin,” n.d.) and the media being water-based, capsaicin aggregated, and the OD (at 600 nm) was unable to be properly measured. In this 48-hour screen, the anti-fungal, Nystatin was used as a positive control for inhibition. Its concentration (in each respective well) was 16 µg/mL, which was shown to
have inhibition against \textit{F. oxysporum} based on plate optimization experiments (unpublished data by Yong Zhang).

\textbf{Figure 9:} Correlation plot between growth inhibition (\%) values of \textit{Fo}32931 (shown with x points) and \textit{Fo}4287 (shown with o points) and the corresponding amount of capsaicin in 50ex pepper extracts (shown by table 5). All samples were corrected for capsaicin that was present in the well.

To put the capsaicin estimation data and growth inhibition data together, a correlation plot was made. The calculated percent growth inhibition (Figure 7 & 8) against each of the three replicates of each of the strains (\textit{Fo}32931 and \textit{Fo}4287) from the 48-hour screen were plotted against the amount of capsaicin in each of the extracts (shown in table 5). Since all treatment volumes were 4 µL inside a total well volume of 200 µL, all estimated capsaicin concentrations were adjusted to 2\% of the amount in the tested (through HPLC) 1 mL of 50ex. Since the 100ex extracts were not run through HPLC, their inhibition values were unused.
**Figure 10:** Plot showing growth curves for the 48 hours screen against *Fo32931*, with a focus on the two MSC extracts. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25 °C and a spore concentration of ~1.26E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested.

**Figure 11:** Plot showing growth curves for the 48 hours screen against *Fo32931*, with a focus on the two MSC extracts. The color key for each line is in the top left of the figure.
Screen conditions included: an incubation temperature of 25 °C and a spore concentration of ~1.26E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested.

Specifically, MSC50ex showed the strongest inhibition out of all the extracts against both Fo32931 and Fo4287 (Figure 11 and Figure 12 respectively). The two WC extracts which had no significant effect against either of the Fusarium strains (Figure 23 and Figure 24 for Fo32931 and Fo4287 respectively). In the screen against Fo32931, CR100ex and CR50ex were close behind the WC extracts in their inability to inhibit growth in addition to having the largest standard deviations in the set (Figure 25). While not very significant, CR100ex delayed growth of Fo4287 (Figure 26). The CRC100ex extract had weaker inhibition than MSC extracts in both screens (Figure 27 and Figure 28 Fo32931 and Fo4287, respectively), but the effect of cRC100ex extract on growth reduction was on average stronger than that of WC extracts and CR extracts. Though there were previous difficulties with making accurate measurements on capsaicin due to its inability to dissolve into aqueous solution, a concentration of 0.25 mg/mL allowed for a clearer read. In the case of Nystatin and 95% capsaicin (0.25 mg/mL), negative values for area under the curves are, due to the removal of background (first three points averaged and removed) and should be recognized as little to no growth. For both Fo32931 and Fo4287, capsaicin significantly reduced growth (Figure 12 and Figure 13, respectively).
Figure 12: Plot showing growth curves for the 48 hours screen against Fo32931, with a focus on growth with 0.25 mg/mL capsaicin. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25 °C and a spore concentration of ~1.26E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested.
Figure 13: Plot showing growth curves for the 48 hours screen against Fo4287, with a focus on growth with 0.25 mg/mL capsaicin. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25 °C and a spore concentration of ~1.25E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested.

Out of all the tested pepper cultivar extracts, MSC50ex is the strongest against all the tested Fusarium isolates and also had the highest level of capsaicin. If the estimation of capsaicin in extracts (Table 5) from the line of best fit (Figure 6) is accurate, the trend in reduced/delayed growth follows the increase of capsaicin concentration. This trend is shown for each of the examined isolates in Figures 29 and 30 for Fo32931 and Fo4287 respectively. The correlation for the line of best fit is strong with a R² values of 0.85 and 0.92 for Fo32931 and Fo4287 respectively.

Taken together, the OD experiments show that the capsaicin-containing pepper extracts have varying levels of inhibition if any at all (in the case of WC). With more material, further purification could have been done and would have likely resulted in a stronger extract. In addition, with more extract, another series of HPLC tests could have been conducted with the stronger inhibiting 100ex samples. To find a more direct correlation and between capsaicin concentration and growth inhibition, capsaicin concentrations that
matched the amount of capsaicin in extracts could have also been used in optical density screens after the HPLC. Doing these optical density screens could have answered if there was some type of synergy occurring in the extracts that was increasing inhibition (if the equivalent capsaicin standard showed less inhibition. These experiments show that *F. oxysporum* is an excellent model for also testing plant secondary metabolites against. Due to its wide host range, there are many isolates that can be tested against. With *F. oxysporum*’s relevance in damaging agricultural products, anything that can severely inhibit its growth is notable for further investigation.
Figure 14: Overview figure with 50 subplots showing the expression pattern of each mapped TE for Fo5176 during the infection course. The vertical axis of each subplot changes to stay within the range of the normalized expression. X-axis is measured in hours post inoculation. Plots with a single point means the isolate did not have the TE or mapped expression reads across all time points. TEs that had similar patterns are listed in table 1.
Figure 15: Overview figure with 50 subplots showing the expression pattern of each mapped TE for Fo47 during the infection course. The vertical axis of each subplot changes to stay within the range of the normalized expression. X-axis is measured in hours post inoculation. Plots with a single point means the isolate did not have the TE or mapped expression reads across all time points. TEs that had similar patterns are listed in table 2.
**Figure 16:** Vertical axis bar is not consistent across all samples. Y-axis is measured in milli-absorption units (mAU) and x-axis is measured in minutes retention time. Extract sample is listed on the left. WC = Habanero White Chili, CR= Carolina Reaper Chili, MSC = Moruga Scorpion Chili, CRC = Carolina Reaper Chili. These chromatograms are from 1mL of unfiltered (u) liquid extract from the designated cultivar that was evaporated. Before being run in the HPLC, 1mL of ethanol (100%) was added and the solution mixed. Arrows designates capsaicin associated peak. As shown by chromatograms of standard capsaicin, (figure 4) capsaicin is responsible for peak at ~6.6 min retention time.
Figure 17: This plot is an extension of the linear line of best fit referring to the capsaicin standard and the associated calibration curve. Black line connects each of the serial dilution area values. Dotted red line is linear line of best fit. Maximum capsaicin standard was 1 mg/mL capsaicin (0.95 mg/mL actual). Equation for line of best fit is on right side of the figure with an R² value of 0.9983.

Figure 18: Subplots showing each of the growth curves for the 48 hours screen against Fo32931. The spore concentration of each well (with culture) was ~1.26E6 spores/mL. Absorbance was measure at 600 nm. Plate was incubated at 25 °C, and was covered. All plots have an x-axis measured in time (n = 30 minutes) and a y-axis with OD (absorbance). Each subplot has their condition labeled above their plot. Each line represents one replicate. Controls reside in the three bottom right panels.
**Figure 19:** Subplots showing each of the growth curves for the 48 hours screen against *Fo4287*. The spore concentration of each well (with culture) was ~1.25E6 spores/mL. Absorbance was measured at 600 nm. Plate was incubated at 25 °C. and was covered. All plots have an x-axis measured in time (n = 30 minutes) and a y-axis with OD (absorbance). Each subplot has their condition labeled above their plot. Each line represents one replicate. Controls reside in the three bottom right panels.
Figure 20: Overview for the 48 hours screen against *Fo32931*. Bar graph showing the total area (y-axis) under each of the growth curves when in the presence of each pepper extract or control (x-axis). The spore concentration of each well (with culture) was ~1.26E6 spores/mL. Plate was incubated at 25 °C. and was covered. Controls: Normal growth, ethanol, and Nystatin are the left-most bars.
Figure 21: Overview for the 48 hours screen against Fo4287. Bar graph showing the total area (y-axis) under each of the growth curves when in the presence of each pepper extract or control (x-axis). The spore concentration of each well (with culture) was \( \sim 1.25 \times 10^6 \) spores/mL. Plate was incubated at 25 °C. and was covered. Controls: Normal growth, ethanol, and Nystatin are the left-most bars.

Figure 22: Plot representing the effect of trying to do an optical density reading with wells containing capsaicin. Spore concentration in each respective well was \( \sim 1.8 \times 10^6 \) spores/mL. The x-axis is time in hours and the y-axis is a measure of the absorbance.
Figure 23: Plot showing growth curves for the 48 hours screen against Fo32931, with a focus on the two WC extracts. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25°C and a spore concentration of ~1.26E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested.

Figure 24: Plot showing growth curves for the 48 hours screen against Fo4287, with a focus on the two WC extracts. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25°C and a spore concentration
of ~1.25E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested.

Figure 25: Plot showing growth curves for the 48 hours screen against Fo32931, with a focus on the two CR extracts. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25°C and a spore concentration of ~1.26E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested.
Figure 26: Plot showing growth curves for the 48 hours screen against *Fo4287*, with a focus on the two CR extracts. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25°C and a spore concentration of ~1.25E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested.

Figure 27: Plot showing growth curves for the 48 hours screen against *Fo32931*, with a focus on the two CRC extracts. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25°C and a spore concentration
of ~1.26E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested.

**Figure 28:** Plot showing growth curves for the 48 hours screen against *Fo4287*, with a focus on the two CRC extracts. The color key for each line is in the top left of the figure. Screen conditions included: an incubation temperature of 25°C and a spore concentration of ~1.25E6 spores/mL. The plate was covered and was read at 600 nm (absorbance) every 30 minutes. Error bars represent the standard deviation of the three replicate wells tested.

**Figure 29:** This figure represents the line of best fit for the correlation between the replicates of growth inhibition for *Fo32931* with each of the extracts (shown in Figure 9 with “x”) and the corresponding concentration of capsaicin.
Figure 30: This figure represents the line of best fit for the correlation between the replicates of growth inhibition for *Fo4287* with each of the extracts (shown in Figure 9 with “x”) and the corresponding concentration of capsaicin.

General model:
\[ f(x) = \frac{a}{1+\exp(-b\times x)+c} \]

Coefficients (with 95% confidence bounds):
\[ a = 6.699 (-0.4369, 13.83) \]
\[ b = 31.95 (17.38, 40.51) \]
\[ c = -0.9269 (-1.004, -0.8493) \]

Goodness of fit:

- SSE: 1268
- R-square: 0.9284
- Adjusted R-square: 0.9164
- RMSE: 10.28
BIBLIOGRAPHY


