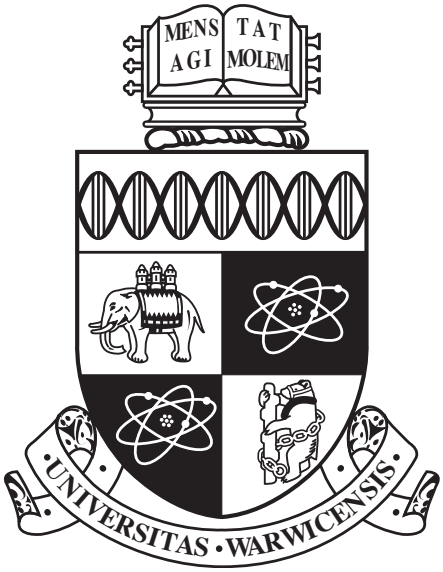
Examining identity, phylogeny and pathogenicity factors in *Fusarium* species affecting pea

Sascha Jenkins

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Life Sciences



University of Warwick, School of Life Sciences

September 2018

Table of contents

List of Figures

List of Tables

Abbreviations

ALG	Algeria
ANOVA	Analysis of variance
Avr	Avirulence
BLAST	Basic Local Alignment Search Tool
bp	Base pair
CAZy	Carbohydrate-active enzymes database
CAZymes	Carbohydrate active enzyme
CDC	Conditionally dispensable chromosomes
cDNA	Complementary DNA
CR	Czech Republic
cv	Cultivar
d.f.	Degrees of freedom
DE	Differential expression
DEG	Differentially expressed genes
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic acid
dpi	Days post inoculation
DSP	Darkskin perfection
ER	Endoplasmic reticulum
ETI	Effector triggered immunity
F	Forward
F. oxysporum	*Fusarium oxysporum*
f. sp.	*Forma specialis*
f. spp.	*Formae speciales*
FOC	*F. oxysporum* f. sp. *cepae*
FOL	*F. oxysporum* f. sp. *lycopersici*
FOP	*F. oxysporum* f. sp. *pisi*
FPKM	Fragments per kilobase million
FTF	*Fusarium* transcription factor
g	Gravitational force
GLM	General linear model
HGT	Horizontal gene transfer
hpi	Hours post inoculation
HR	Hypersensitive response
HRI	Horticulture Research Institute
I protein	Immunity proteins
ITS	Internal transcribed spacer region
kb	Kilobase
LS	Lineage specific
LSD	Least significant difference
MAMP	Microbe associated molecular patterns

MEGA	Molecular Evolutionary Genetics Analysis
mimp	Miniature impala
MITE	Miniature inverted repeat transposable element
mtSSU	*Mitochondrial small subunit*
NB-LRR	Nucleotide-binding leucine rich repeat
NCBI	National Centre of Biotechnology Information
NGS	Next generation Sequencing
ORF	Open reading frame
p	Probability
PAMP	Pathogen associated molecular patterns
PCA	Principal component analysis
PCR	Polymerase chain reaction
PDA	Potato dextrose agar
PDB	Potato dextrose broth
PGRO	Processors and growers research organisation
PHI-base	Pathogen-Host Interaction database
PR	Pathogenesis related
PRR	Pattern recognition receptors
qPCR	Quantitative polymerase chain reaction
R	Reverse
R genes	Resistance genes
RBP	*RNA polymerase largest subunit*
rDNA	Ribosomal deoxyribonucleic acid
REML	Restricted or residual maximum likelihood
RNA	Ribonucleic acid
RNAseq	RNA sequencing
rpm	Revolutions per minute
RR	Root rot
SDW	Sterile distilled Water
SE	Standard error
SIX	*Secreted In Xylem*
TE	Transposable element
TE buffer	Tris-Ethylenediaminetetraacetic acid (EDTA) buffer
TEF	*Translation elongation factor 1a*
TMD	Transmembrane domain
TUB2	*β–tubulin*
UK	United Kingdom
UN	Unknown
USA	United States of America
USER-Brick	Uracil specific excision reagent
VCG	Vegetative compatibility group

Acknowledgements

Firstly, I would like to thank my supervisor Dr John Clarkson for giving me the opportunity to complete this PhD. After completing a summer project and my undergraduate project in your group, you allowed me to return yet again to undertake this PhD project, for which I am very grateful. I have learnt so much from all these great opportunities and gained so much more confidence and belief in myself, although I don't always show it. Thank you for your enthusiasm, support and guidance (and for all the time you've spent reading this thesis). I would also like to thank Dr David Chandler for being an amazing personal tutor, for believing in me and for keeping me on track throughout my undergraduate degree. Without you I never would have considered a career in science and your passion for science has definitely rubbed off.

I would like to thank Dr Richard Harrison, my second supervisor, for all the support and guidance throughout the project, particularly the genome analysis side of the project. Without your involvement and ideas, it simply wouldn't have happened. I am also incredibly grateful to Dr Andrew Armitage for teaching me all I know about bioinformatics and genomics. I would not have been able to complete this side of the project without you giving up so much time and putting up with my constant pestering on a daily basis.

I would like to thank Dr Andrew Taylor for all your help and advice over the many years. You taught me so much during my undergraduate projects, and I would not have considered research if you hadn't made them so enjoyable. Your support and advice throughout this PhD has been invaluable and I could not have sampled all those many pea fields without your help!

I am extremely thankful to Alison Jackson for all your support in the lab (and out) throughout this project and for teaching me so many techniques and skills. You always had a time saving tip to share and frequently took time out of your day to show me how to do things, help me set up experiments and troubleshoot every problem I encountered. I am thankful for your friendship, kind words and for looking after me during this whole experience. I would have been completely lost without you.

I would like to thank Andrew Mead for help with statistical analysis, Richard Stark for help with bioinformatics and horticultural services, especially Peter, Trevor and Sally for all the support and assistance they provided with the many glasshouse experiments. I also need to thank the growers who let me sample their fields and my advisory panel members for the useful discussions and insights into the project.

Completing this PhD would have been so much harder without the company of people from Warwick Crop Centre. Nicole, you are such a caring and lovely person, you always lightened my days even if things weren't going so well. Zofia, it has been pleasure sharing this experience with you. I would also like to thank everyone else at Wellesbourne for the support and encouragement, and for making the office a great place to be.

I'm so pleased to have shared this journey, from start to finish, with Kathryn Hales. From the first day when we met at main campus to collect our university cards to the very last day of submission, we have shared every high, every low and all the things in between. It would have been so much harder doing this PhD without you, and far less fun. I will be sad not to see you every day, and spend too much of our days chatting rather than working! You really did keep me sane and are a truly great friend.

Obviously, I would like to thank my funders, BBSRC, for providing financial support and for the opportunity to study for this PhD. I have gained so much from it.

A huge thank you to my parents for your continued support, encouragement and help with everything else in my life. Mandy, you always take care of me no matter what and you have been constantly supportive in everything I have chosen to do and I am so thankful to call you my sister. I genuinely wouldn't be who I am without your love and guidance. Nicky, you have kept me company during the hard days of writing, gave me perspective and made me laugh on countless occasions and I'm so glad to have you in my life.

Finally, I owe my greatest thanks to my partner in crime, for your love, support and encouragement. You have helped me in so many ways throughout this PhD and your enthusiasm for science is infectious. Pratik, I absolutely would not have been able to do this without you, and I thank you for sticking by me through this whole process. Thank you for believing in me even when I didn't and for knowing exactly what to do and say when I couldn't see a way forward. I will be forever grateful.

xx

Declaration

The work contained in this thesis is original research carried out by the author, at the University of Warwick, between October 2014 and September 2018. It has been composed by the author and has not been submitted for any other degree, or at any other institution. Content generated through collaboration, or adaptation of previous work has been fully acknowledged and referenced.

Signed:_____ Date: _____

Sascha Jenkins

Abstract

Fusarium oxysporum is a fungal plant pathogen responsible for causing disease in many economically important crops. There are many 'special forms' (*formae speciales*) which cause infection in specific plant hosts. *F. oxysporum* f. sp. *pisi* (FOP) is the causal agent of wilt in peas and has been reported in every country where peas are grown commercially. Disease is currently controlled using resistant cultivars but due to the single gene resistance mechanism in pea there is a constant risk of resistance breakdown. Therefore, the aim of this work was to understand the molecular and genetic factors affecting pathogenicity of different FOP races. Additionally, the diversity of *Fusarium* species affecting UK peas was assessed.

Isolations from infected peas in the UK indicated that numerous species could potentially be responsible for causing root rot. *F. oxysporum* isolates from infected peas separated into distinct clades compared to FOP isolates, leading to the conclusion that root rot causing isolates are more prevalent in UK pea fields than wilt causing isolates. Pathogenicity assays were developed for assessing root rot (seed inoculation) and wilt pathogens (root dip), which confirmed the difference between these isolates.

Whole genome sequencing of three pathogenic FOP isolates revealed the presence of *Secreted In Xylem* (*SIX*) genes, and their confirmation in multiple isolates of races 1, 2 and 5 revealed a potential race specific complement of *SIX* genes in FOP. Significant upregulation of these genes *in planta* was observed over the time course of infection. Additional putative effectors predicted by genome analysis were also shown to be upregulated *in planta* at 96 hpi using RNAseq.

Finally, FOP race type was confirmed, with results suggesting that putative race 5 isolates were most likely a subset of race 2, and therefore only clear differences between race 1 and race 2 could be established with effector gene profiles.

Overall, this work shows that there are multiple factors influencing pathogenicity of FOP towards pea. The race specific *SIX* gene complements, and other novel race specific effectors, indicate targets for molecular identification of FOP races.

1. Introduction

1.1 The value of vining peas and their production

Pea (*Pisum sativum L.*) belongs to the Leguminosae family and is an important legume grown all over the world, for both human and animal consumption. Vining peas (field/green pea) and combining peas (dry pea) are the two major pea crops grown for fresh vegetable/frozen (vining) and the dried pulse crop (combining) markets (Biddle & Cattlin, 2007). In 2016, nearly 20 million tonnes (Mt) of green peas were produced from approximately 2.6 million hectares worldwide, and just over 14 Mt of dry peas were produced from approximately 7.6 million hectares (FAOSTAT, 2016). The largest producers of field pea are China (approx. 12 Mt) and India (approx. 4 Mt), followed by the USA and France (31,000 and 23,000 tonnes respectively). This is in comparison with dry pea where Canada and Russia are the two largest producers (4.6 and 2.1 Mt respectively). The UK was the sixth largest producer of vining pea in 2016 with approximately 150,000 tonnes produced from 35,000 hectares, with a value of £154 million to the UK economy (FAOSTAT, 2016).

The pea is a cool season, herbaceous, trailing annual plant and is usually one of the first crops to be sown in spring (Kraft & Pfleger, 2001). Exact sowing time is dependent on cultivar, land and the fate of the peas; for example, processing factories for vining pea will require a steady supply of peas and therefore sowings are from early spring (March) up until the end of May (Maguire, 2014). The growing season for vining peas is quite short at around 10-12 weeks (Moore *et al.*, 2018), depending on climate variations, and harvest time is decided based on the maturity of the peas determined by using a pea tenderometer. This measures maturity through the force required to shear a sample of peas, and the readings and yield from this sample can also determine grower payments (PGRO, 2015). Vining pea production is located on the eastern side of the UK, ranging from East Anglia to north of Montrose in Scotland (PGRO, 2015), as the maritime climate provides optimal growing conditions, but the area is also limited by the requirement for fields to be close to processing factories (Moore *et al.*, 2018). This is because many peas picked for freezing need to be frozen within 150 minutes to retain the freshness and taste in the finished product (Green & Foster, 2005). Pea plants grow best on fertile, light, well-drained and humus-rich soils (Tulbek *et al.*, 2017) but are particularly susceptible to soil

compaction (Vocanson *et al.*, 2006). Peas are harvested throughout the world using 'complete' harvesters, which combine the operations of picking, threshing and cleaning in one step. Pods are stripped from the vine, passed through a threshing drum where the peas are removed from the pods, where they fall through a set of screens and are cleaned to remove the waste plant material which is returned to the field (PGRO, 2015). Peas are transported to one of the eight freezing factories or the one canning factory currently available in the UK (PGRO, 2015).

1.2 The importance of peas

The bulk of the mature pea seed is made up of two large fleshy cotyledons, in addition to stored reserves that sustain early seedling growth (Kraft & Pfleger, 2001). At germination, the cotyledons remain underground until emergence when the epicotyl penetrates the soil surface (Figure 1.1). The first two nodes above the cotyledon usually contain rudimentary leaves which are followed by the first true leaves and then a succession of leaves arising from defined nodes along the stem (Kraft & Pfleger, 2001). Peas are self-pollinating, shedding pollen in the closed flower, so fertilisation occurs around 1 day before it opens (Kraft & Pfleger, 2001). The pea pods pass through a flat and round pod stage, whereby the pod lengthens and widens soon after fertilisation (flat pod) and starts to become round as the peas inside increase in size and weight though cell enlargement. Pod maturity can take between 24 – 30 days, and after this time will enter the dry seed stage (Kraft & Pfleger, 2001) where the seeds and pods begin to lose water and dry out.

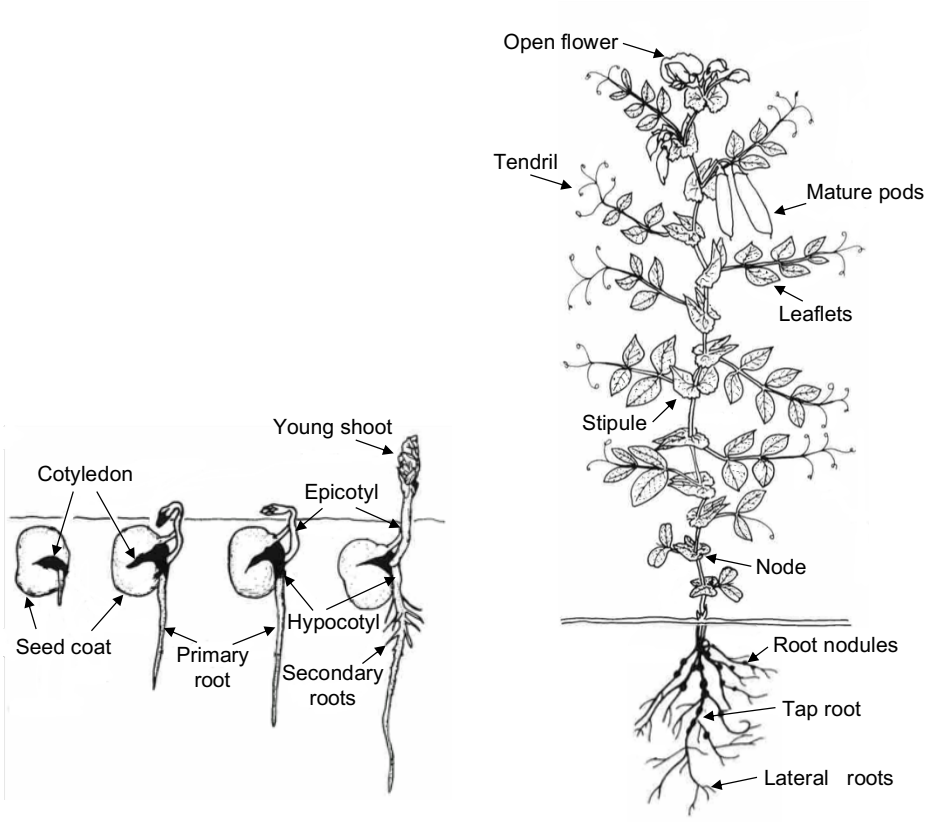


Figure 1.1 Germinating pea seed (left) and mature pea plant in flower and pod setting stage (right). Adapted from Kraft and Pfleger (2001).

As peas are legumes they are able to form symbiotic relationships with nitrogen fixing bacteria called rhizobia. The low O_2 environment in root nodules allows rhizobia to convert atmospheric unreactive N_2 to ammonia, using the enzyme nitrogenase, which can then be utilised by plants (Bhattacharjee *et al.*, 2008). Although legumes utilise a lot of the fixed nitrogen from rhizobia, some residual nitrogen will remain in the soil that can be utilised by following crops and it is often observed that yields are greater when legumes are included in the rotation (Peoples *et al.*, 2009). Peas and beans are usually included in arable rotations, as well as cereals and wheat, and due to the high nitrogen requirements in some of these crops, legumes can help to reduce the economic inputs by reducing nitrogen requirements (Biddle & Cattlin, 2007). They are favoured as break crops in cereal rotations, as the ability to sow in the spring allows the land to be fallow for the winter and provides opportunities to control perennial weeds which is often difficult in a cereal rotation (Biddle & Cattlin, 2007). However, it is recommended that peas are not included in rotation more than once in 5-7 years to reduce the build-up of soil borne pests and diseases (PGRO, 2015).

1.2.1 Nutritional benefits of pea

Peas are recognised as an inexpensive, readily available source of protein, complex carbohydrates, vitamins and minerals and are low in fat (Dahl *et al.*, 2012). The high fibre content of peas has been shown to reduce the glucose response after consuming pea flour instead of whole-wheat flour, which may be beneficial in managing type two diabetes (Marinangeli *et al.*, 2009). Fibre rich diets containing peas have also been shown to lower blood pressure and improve gastrointestinal function (Dahl *et al.*, 2012). They are a good source of vitamin A, vitamin C, folate, thiamine (B1), iron and phosphorus (yes-peas!, 2018), in addition to phenolic compounds which may help to protect against cancers and inflammatory diseases (Dahl *et al.*, 2012). Protein content of pea cultivars range from 23 – 31 % (N x 5.6), providing an important source of essential dietary protein and an alternative to animal proteins (Boye *et al.*, 2010).

There is growing awareness of the usefulness of peas and other pulses in the diet, which has seen a steady rise in interest of using peas, other pulses and ingredients derived from them in novel food products (Boye *et al.*, 2010). These products include pea protein powder, crisps/baked snacks and pea milk, which all offer healthier alternatives to traditional foods in these categories.

1.2.2 Important diseases of pea

Pea production is threatened by economically important diseases with potential to cause significant losses. The most important diseases of pea include downy mildew (caused by *Peronospora viciae*), leaf and pod spot (caused by three types of fungi: *Ascochyta pisi*, *Mycosphaerella pinodes* and *Phoma medicaginis*), grey mould (caused by *Botrytis cinerea*), sclerotinia disease (*Sclerotinia sclerotiorum*), pea mosaic virus and pathogens of the pea root rot complex (e.g. *Rhizoctonia*, *Pythium*, *Aphanomyces* and *Fusarium* species; discussed in more detail in 1.7.2) (PGRO, 2015). Fungi are the most common cause of diseases in pea (Kraft & Pfleger, 2001) most likely because they are able to spread quickly and across a large area due to numerous types of spores, which can also reside in the soil for a long time (Kraft & Pfleger, 2001). Of particular importance are *Fusarium* species, as numerous species can infect pea, causing root rot and wilt diseases.

1.3 *Fusarium*

1.3.1 *Fusarium* genus

The genus *Fusarium* belongs to the Ascomycota phylum of fungi (Table 1.1) and was first identified by Link in 1809 based on the presence of distinctive banana shaped conidia (Leslie & Summerell, 2006). Phylogenetic studies indicate that *Fusarium* originated ~91.3 million years ago which coincided with the emergence of flowering plants (Smith *et al.*, 2010; Ma *et al.*, 2013). *Fusarium* contains a wide variety of filamentous fungi, many of which are important plant pathogens, but can also cause disease in humans and animals (Leslie & Summerell, 2006; Moretti, 2009). Some of the most economically important species of plant pathogens belong to *Fusarium*, and affect agriculture and horticulture globally (Summerell *et al.*, 2010). There are over 80 species of *Fusarium* currently recognised (Leslie & Summerell, 2006) responsible for vascular wilts, root rots and head blights in many agriculturally important crops worldwide (Nelson *et al.*, 1994). It has been reported that 81 out of 101 economically important plants have at least one associated *Fusarium* disease (Leslie & Summerell, 2006). *Fusarium* species are found in many climatic regions including temperate, tropical, arctic and desert conditions (Nelson *et al.*, 1994). They can survive in almost any biome including forest, grasslands, agricultural and alpine zones, as well aquatic and man-made environments (Balmas *et al.*, 2010). Some of the most important pathogenic *Fusarium* species are responsible for major plant diseases worldwide, for example, *Fusarium* head blight of wheat and other small grain cereals caused mainly by *Fusarium graminearum*, but also *F. culmorum, F. avenaceum* and *F. poae* (Nicholson *et al.*, 2003). As well as being plant pathogens, some *Fusarium* spp. have been found to be non-pathogenic and even beneficial, living as saprophytes and preventing plant diseases (Edel *et al.*, 1997; Alabouvette *et al.*, 2009).

Table 1.1 Taxonomy of *Fusarium* genus with five examples of plant pathogenic species (NCBI, 2018)

***Fusarium* taxonomy**	
Kingdom	Fungi
Phylum	Ascomycota
Class	Sordariomycetes
Order	Hypocreales
Family	Nectriaceae
Genus	*Fusarium*
Species	*F. oxysporum*
	F. solani
	F. graminearum
	F. avenaceum
	F. redolens

1.3.2 Sub - specific ranks to further classify *Fusarium* species

Some morphologically defined *Fusarium* species are complexes containing mostly indistinguishable sub-species (van Diepeningen *et al.*, 2014). Sub-specific ranks such as *forma specialis*, races and vegetative compatibility groups are used to distinguish between isolates of such species. *Forma specialis* (f. sp.), or the plural *formae speciales* (f. spp.), refers to specific strains of a species that only infect one or a small number of host plants and can therefore be differentiated from each other on the basis of pathogenicity (Leslie & Summerell, 2006). *Formae speciales* are not taxonomic entities and are only recognised based on their ability to infect and cause disease in specific plant hosts (Summerell & Leslie, 2011). They are thought to have evolved though convergent evolution and many are therefore polyphyletic (Baayen *et al.*, 2000; Summerell & Leslie, 2011); hence, isolates from one f. spp. may be more closely related to isolates infecting other hosts than to each other (O'Donnell *et al.*, 1998). The most common example is within *F. oxysporum* species, where more than 150 host specific *formae speciales* have been described (Baayen *et al.*, 2000).

Formae speciales can sometimes further be divided into races, which are specialised to infect certain cultivars of a host species (Biddle & Cattlin, 2007). It is thought that these races have evolved from individual isolates that have developed the ability to overcome specific plant resistance genes present in certain cultivars (Takken & Rep, 2010). New races occur when host resistance is overcome by the loss, gain or mutation of effectors previously recognised by the plant, allowing the pathogen to evade detection (Takken & Rep, 2010). *F. oxysporum* can also be separated into vegetative compatibility groups

(VCGs) which contain isolates capable of fusing and linking compatible hyphae (anastomosis) to produce cells with two haploid nuclei (heterokaryon) (Leslie, 1993). Isolates or strains in different VCGs form a heterokaryon that is unbalanced or transitory (Leslie & Summerell, 2006). Each *forma specialis* can contain multiple VCGs, which are given a number which correlates all the members of the VCG to the *forma specialis* they are associated with (Leslie & Summerell, 2006).

1.4 *Fusarium oxysporum*

One of the most widely dispersed *Fusarium* species is *Fusarium oxysporum*. It is the most economically important plant pathogenic species in the *Fusarium* genus as it infects numerous hosts and causes extensive crop losses (Leslie & Summerell, 2006). It was identified as 5th in a list of the top 10 fungal plant pathogens in terms of scientific and economic importance, just behind *F. graminearum* in 4th place (Dean *et al.*, 2012). *F. oxysporum* is responsible for a vast range of plant diseases, usually causing vascular wilt but also root and bulb losses in many plants (Leslie & Summerell, 2006). Of the numerous f. spp. identified in *F. oxysporum* there are many well studied and economically important examples including *F. oxysporum* f. sp. *cubense* (banana), f. sp. *lycopersici* (tomato), f. sp. *vasinfectum* (cotton), f. sp. *cepae* (onion), f. sp. *phaseoli* (bean), f. sp. *ciceris* (chickpea), f. sp. *lactucae* (lettuce) and f. sp. *pisi* (pea). *F. oxysporum* f. sp. *lycopersici*, *cubense* and *pisi* are examples of f. spp. which contain multiple economically damaging races, for example, f. sp. *cubense* tropical race 4 which is devasting the popular Cavendish cultivar in the tropics (Ploetz, 2015).

1.5 Non-pathogenic *Fusarium* species

As well as many pathogenic species, some *Fusarium* species have been found to be non-pathogenic and can naturally suppress some Fusarium wilts (Edel *et al.*, 1997). This has led to the identification of non-pathogenic isolates as biocontrol agents, functioning by colonising the root surface and preventing pathogens access to infection sites (Alabouvette *et al.*, 2009). One particular isolate, Fo47 has been shown to directly induce resistance in tomato against *F. oxysporum* f. sp. *lycopersici* (FOL) (Fuchs *et al.*, 1997).

1.6 *F. oxysporum* life cycle

F. oxysporum produces three spore types: macroconidia, microconidia and chlamydospores (Figure 1.2), containing nuclei derived from hyphae produced by mitosis (Gordon, 2017). There is no known sexual stage in *F. oxysporum* and to date there have been no reports of structures such as perithecia, asci and ascospores (Gordon, 2017). However, fewer than 20% of *Fusaria* have a known sexual cycle (Ma *et al.*, 2013). Macroconidia are long canoe shaped septate spores that are produced in cushion-shaped aggregates of conidiophores called sporodochia or on conidiophores in the aerial mycelium. Microconidia (Figure 1.3) are small asexual spores produced in mycelium from conidiophores (spore forming structures). Chlamydospores are produced within or on hyphae or macroconidia (Ma *et al.*, 2013).

F. oxysporum can survive extended periods in the absence of a host, due to thick walled chlamydospores which can result in an area remaining infested (Di Pietro *et al.*, 2003). The proximity of roots induces dormant spores to germinate, causing hyphae to adhere to host roots (Di Pietro *et al.*, 2001) and penetrate the root either inter- or intra-cellularly (Rodriguezgalvez & Mendgen, 1995). The mycelium progresses towards the xylem vessels inter-cellularly through the root cortex, and enters the vessels through the pits (Figure 1.2). It then remains in the xylem vessels and uses them to rapidly colonise the host (Di Pietro *et al.*, 2003). Production of microconidia allows *F. oxysporum* to spread quickly around the plant as they detach from the mycelium and are carried upwards in the xylem sap where they germinate and produce more mycelium which penetrates the upper walls of xylem producing more spores in the next vessel (Di Pietro *et al.*, 2003). Wilt symptoms appear as severe water stress due to the accumulation of mycelium blocking the vessels, the production of toxins and from plant defence responses. Once the plant dies, the fungus can invade other plant tissues and sporulate on the plant surface (Di Pietro *et al.*, 2003).

F. oxysporum is spread through movement of contaminated material (either soil or plant debris) by water, wind, and people. Infected plant material and seeds can contribute to long distance dissemination (Kraft & Pfleger, 2001). Infected fields should be treated with containment procedures to ensure spread does not occur between fields.

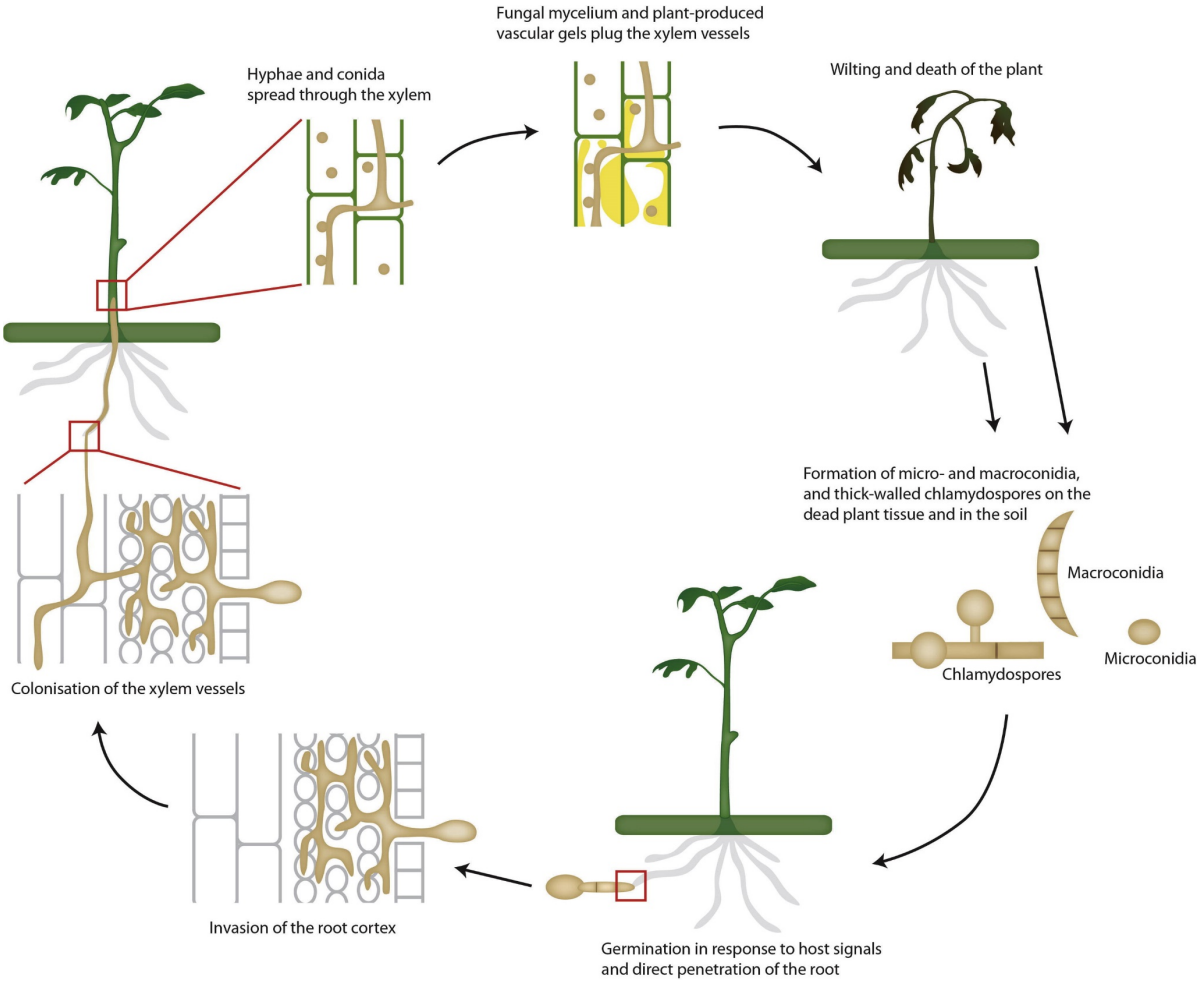


Figure 1.2 Life cycle of *Fusarium oxysporum* (Perez-Nadales *et al.*, 2014)

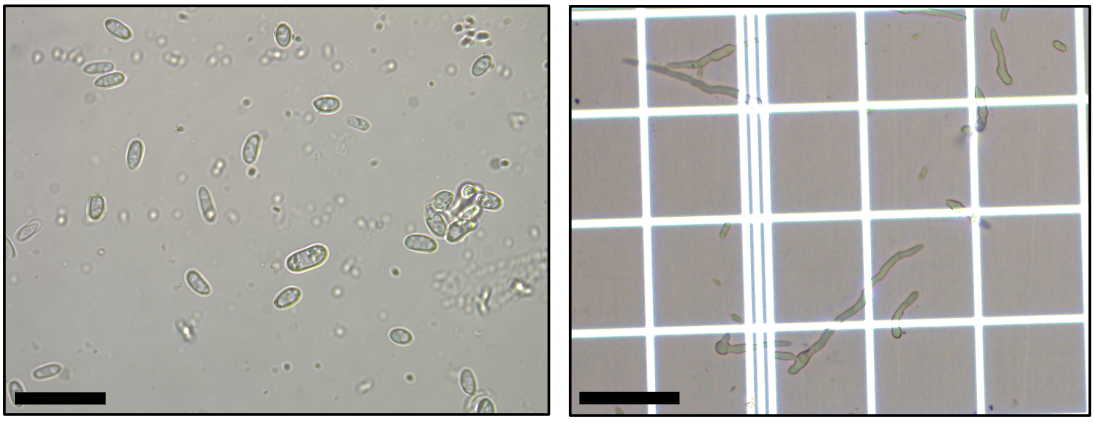


Figure 1.3 Microconidia (left) and germinated microconidia (right) from *Fusarium oxysporum* f. sp. *pisi*. Scale bar: 20 µM (left), 50 µM (right)

1.7 *Fusarium* diseases in pea

1.7.1 *Fusarium oxysporum* f. sp. *pisi* causing pea wilt

Fusarium wilt caused by *F. oxysporum* f. sp. *pisi* (FOP) is a major disease of pea and has been reported in every country where peas are grown commercially (Kraft, 1994). It can result in severe crop losses especially when peas are included in crop rotations more often than recommended, as it allows inoculum levels to build and chlamydospores can remain viable in the soil for 10 years (Kraft *et al.*, 1994). FOP race 1 was identified in 1924 by Jones and Linford in Wisconsin, USA (Jones & Linford, 1925) and by 1928 it had spread to most of the major pea growing areas in the USA (Kraft & Pfleger, 2001). After the introduction of pea cultivars resistant to race 1, a new race emerged which overcame this source of resistance which was named FOP race 2. Additional races 3 and 4 were described in the Netherlands and Canada respectively, but were later re-classified as more virulent forms of race 2 (Kraft & Haglund, 1978). Race 5 was described in 1970, based on the observation that all cultivars known to be resistant to races 1 and 2 were susceptible (Haglund & Kraft, 1970). A further study classified isolates into 11 races, including the previously described races 1, 2, 3, 4 and 5, based on virulence towards different pea cultivars (Armstrong & Armstrong, 1974). However, in 1977 these races were re-evaluated using pea differential cultivars and only the original races of 1, 2 and 5 were considered to be valid (Kraft & Haglund, 1978). It was concluded that most of the isolates grouped into the 11 races by Armstrong and Armstrong (1974) were variants of races 1 and 2; race 5 however is still considered to be a distinct race as determined by Haglund and Kraft (1970). Another new race was identified in 1979, and was classified as race 6 based on its prevalence and pathogenicity to all cultivars resistant to races 1, 2 and 5 (Haglund & Kraft, 1979). FOP races 1 and 2 are currently the most important globally, and are the only races identified in the UK (Biddle & Cattlin, 2007), whereas races 5 and 6 are largely found in Western Washington State, which is becoming less of a concern with resistant varieties and as pea production moves out of the area (Infantino *et al.*, 2006).

The definitive method to determine FOP race is through the host response observed following inoculation of a set of pea differential cultivars. These cultivars have known dominant genes for resistance and have been used in several studies to distinguish

between the four races (Table 1.2) (Kraft & Pfleger, 2001; Bani *et al.*, 2012; Haglund & Kraft, 1979; McPhee *et al.*, 1999). Currently, only four single, separate, dominant genes are known to confer resistance, one gene for each individual race, and so far these are the only genes that have been used to develop FOP resistant pea cultivars (Kraft & Pfleger, 2001).

1.7.1.1 Symptoms of Fusarium wilt in pea

Symptoms caused by FOP begin with the downward curling of the pea leaves (Figure 1.4b) and as the disease develops the leaves turn yellow progressively from the base of the stem to the apex of the plant (Kraft, 1994). Roots initially appear to have no symptoms on the outside but can contain orange – red discolouration in the vascular tissue caused by the pathogen (Kraft, 1994), which extends into the basal part of the stem. Disease can develop rapidly and the entire aerial part of the plant wilts and dies (Figure 1.4b), usually well before pod setting stage in races 1, 5 and 6 (Kraft & Pfleger, 2001). If the roots are free from invading root rot organisms then auxiliary buds may develop from the basal nodes of the plant but their development is limited because of the eventual death of the root system caused by FOP (Figure 1.4d) (Kraft & Pfleger, 2001). FOP races 1, 2 and 5 can affect extensive areas in the field where large patches of yellowing and wilted plants are commonly observed (Figure 1.4a) (Biddle & Cattlin, 2007). However, FOP race 2 has been observed to cause symptoms on individual plants scattered throughout the field, and it more commonly affects plants at the flowering and pod setting stages (Biddle & Cattlin, 2007). Symptoms caused by FOP race 2 are similar to the other races but wilt can develop unilaterally up the stem, causing the leaves on one side to wilt before the other (Figure 1.4c) (Kraft & Pfleger, 2001).

Figure 1.4 Symptoms of Fusarium wilt caused by *Fusarium oxysporum* f. sp. *pisi* (FOP), a) extensive patches of yellowing / wilt in a field of peas, b) downward curling and wilting of leaves on an individual pea plant, c) near wilt symptoms caused by FOP race 2, and d) regrowth occurring at the base of a FOP infected plant.

1.7.2 Role of *Fusarium* species in the root rot complex

The root rot complex in pea comprises of diseases caused by single or combinations of pathogens including *Alternaria alternata*, *Aphanomyces euteiches*, FOP, *F. solani* f. sp. *pisi*, *Pythium* species, *Rhizoctonia solani* and *Didymella pinodella* (Xue, 2003). These pathogens individually or synergistically cause symptoms of seed decay, root rot, seedling blight and wilt (Xue, 2003), which can all reduce crop yield and quality (PGRO, 2015). In the UK, the main pathogens causing root rot are *Fusarium* species, *D. pinodella* and *A. euteiches* (Dr L. Herold, PGRO, personal communication). *Aphanomyces* causes

severe stunting and root decay and often occurs in patches in the field, usually in waterlogged areas (Biddle & Cattlin, 2007). *D. pinodella* can cause yield loss up to 70% and is known to be less specialised than other *Didymella* species, allowing inoculum levels to build when other hosts are incorporated in crop rotations (Barilli *et al.*, 2016). All major UK root rot pathogens can survive for long periods in the soil due to the production of thick walled chlamydospores (*Fusarium*, *Didymella*) and oospores (*Aphanomyces*) (Biddle & Cattlin, 2007).

Fusarium root rot has traditionally been associated with *F. solani* f. sp. *pisi*, which was first described in both the USA and Europe between 1918 – 1923 and can cause up to 30% crop losses in infested fields (Kraft & Pfleger, 2001). Symptoms include reddish brown to black discolouration of roots, stunted growth, yellowing of leaves and stems and necrosis of the basal foliage (Figure 1.5) (Biddle & Cattlin, 2007). Root rot and wilt can be difficult to distinguish in the field due to similar symptoms, especially in the late stages of infection (Biddle & Cattlin, 2007). Several other *Fusarium* species such as *F. avenaceum* and *F. oxysporum* have also been isolated from infected pea roots and may also be important components of the root rot complex in addition to *F. solani* in North Dakota (Chittem *et al.*, 2015). In pathogenicity tests, all these species were capable of causing root rot in pea but to different degrees; interestingly *F. avenaceum* caused more severe symptoms than *F. solani* in this study (Chittem *et al.*, 2015). The observation that several different *Fusarium* spp. can cause pea root rot and that *F. avenaceum* can be particularly virulent has also been reported in other areas including Canada (Feng *et al.*, 2010) and Sweden (Persson *et al.*, 1997). Overall, root rot in pea is caused by a complex of pathogens, which makes it particularly difficult to manage and control.

Figure 1.5 Fusarium root rot symptoms causing stunted and yellow plants in the field (top) and the classic black discolouration of the upper root (bottom).

1.7.2.1 Control of FOP and Fusarium root rot in pea

The most effective method for controlling FOP is through resistant cultivars, however, the requirement for good cultural practices and crop rotation is still a necessity. Resistance to the four races of FOP is conferred by single dominant genes, which have been successfully introgressed into pea cultivars from wild relatives (Bani *et al.*, 2012; Infantino *et al.*, 2006). However, due to this monogenic resistance there is a high risk of resistance breakdown as it could easily be overcome by the appearance of new pathogen

variants. There is therefore a need to continue to identify sources of novel resistance, with an emphasis on polygenic mechanisms (Bani *et al.*, 2012).

There are currently no commercial pea cultivars resistant to Fusarium root rot pathogens, but some with increased tolerance have been developed (Biddle & Cattlin, 2007). FOP resistant cultivars have no benefit against Fusarium root rot pathogens, including *F. oxysporum* root rot isolates. Good tillage practices that prevent or reduce soil compaction and promote favourable soil moisture, allowing roots to develop well, can help to reduce root rot (Kraft & Pfleger, 2001). Crop rotation is however, the most important factor in reducing the inoculum level in the soil and prevent it building to damaging levels over time and it is recommended that peas only appear once in a 5 – 7 year rotation (PGRO, 2015). *Fusarium* diseases are difficult to control with fungicides, as the pathogen is soil borne and may not come into contact with the chemical (Sharma *et al.*, 2010).

1.8 Pathogenicity tests for *Fusarium* species affecting pea

Currently, disease assays are the only certain way of discriminating host range and race of *F. oxysporum* (van Dam *et al.*, 2016). The universal assay for determining the pathogenicity and race type of FOP involves inoculation of trimmed roots by dipping in spore suspensions, using a set of pea cultivars with known dominant resistance genes for the different races (Table 1.2) (Kraft, 1994; Kraft & Pfleger, 2001). Pea differential cultivars should be maintained and tested for homozygosity of resistance genes to different races to ensure these cultivars have the predicted reaction during isolate race typing (Kraft & Pfleger, 2001). Alternative assays are used for testing Fusarium root rot in pea, which generally involve inoculation of the soil or seed before sowing (Kraft *et al.*, 1994). These different inoculation assays allow for the development of symptoms for wilt or root rot (Figure 1.6) in an accelerated time frame than would occur naturally in a field environment, allowing virulence of the pathogen and resistance in plants to be tested more quickly. Even with the recent advent of molecular techniques for distinguishing between f. spp. and races (Lievens *et al.*, 2009), pathogenicity tests remain the only certain way to determine the full host range and pathogenicity of isolates.

Table 1.2 Pea differential cultivars and their reaction to inoculation with races 1, 2, 5 and 6 of *Fusarium oxysporum* f. sp. *pisi* (FOP) (Adapted from Kraft & Pfleger, 2001)

Differential cultivar	**Races of *F. oxysporum* f. sp. *pisi***			
	1	2	5	6
Little Marvel	S	S	S	S
Darkskin Perfection (DSP)	R	S	S	S
Mini	S	R	S	S
New Era	R	R	S	S
Sundance II	R	S	R	S
Grant	R	S	S	R
New Season	R	R*	S	S
WSU 23	R	R	R	S
WSU 28	R	S	R	R
WSU 31	R	R	R	R

S = Susceptible, R = Resistant, * Reaction may vary with isolate

Figure 1.6 Pathogenicity tests showing the use of a seed inoculation assay (top) and root dip method in individual pots (bottom) to test root rot and wilt causing *Fusarium* isolates (respectively) in pea

1.9 Molecular identification of *Fusarium* species

Morphological characteristics are still one of the most commonly used criteria for identifying *Fusarium* species. However, suitable media and correct conditions are required, and even then, practice and experience is needed to ensure accurate identification (Leslie & Summerell, 2006). Differences in the shape of macroconidia are central to the identification of many *Fusarium* species as well as other spores such as microconidia and chlamydospores. Although these characteristics are limited in number, they can vary depending on environmental conditions, and subtle differences between species can easily be misinterpreted (Leslie *et al.*, 2001). In addition, there are too few well-defined morphological characters for the number of species that need to be distinguished (Leslie *et al.*, 2001). Morphology also does not allow distinctions to be made between f. spp., and therefore these are still difficult to identify without pathogenicity tests (Leslie & Summerell, 2006). To overcome these problems, molecular identification using gene sequencing and phylogenetic analysis has now become the most common way to distinguish species due to its increased accuracy and reliability. The internal transcribed spacer region (*ITS*) of the rDNA has been proposed as the universal bar code region for fungi (Schoch *et al.*, 2012) but it is not suitable for *Fusarium* as two non-orthologous sequence types can be isolated from within a single biological species and from single isolates (O'Donnell & Cigelnik, 1997; Leslie *et al.*, 2001). Instead, other housekeeping genes such as *translation elongation factor 1α* (*TEF*), *RNA polymerase largest subunit* (*RPB1*) and *RNA polymerase second largest subunit* (*RPB2*) are commonly used as they are single copy and can be amplified by PCR and sequenced using conserved primers (O'Donnell *et al.*, 2010). Sequencing additional housekeeping genes such as *mtSSU* (*mitochondrial small subunit*) and *β-tubulin* can result in greater resolution in molecular phylogenies of *Fusarium* and higher bootstrap support (O'Donnell & Cigelnik, 1997). Online libraries of *TEF*, *RPB1* and *RPB2* sequences for many *Fusarium* species exist, helping users identify sequences by conducting BLAST searches against the database (Geiser *et al.*, 2004; O'Donnell *et al.*, 2010), improving the accuracy of species identification. However, molecular identification using housekeeping genes does not distinguish between *F. oxysporum* f. spp., which still requires the use of pathogenicity tests (Leslie & Summerell, 2006). In recent years, effector genes unique to f. spp. or distinct in sequence identity between f. spp. have been used as a new molecular approach

for identification (Lievens *et al.*, 2009; Fraser-Smith *et al.*, 2014). This is discussed in more detail below (Section 1.10.2).

1.10 Plant – pathogen interactions

There are various forms of plant-microbe interactions such as commensalism (microbes living off natural plant exudates), endophytism (microbes living within a host, without affecting its fitness), symbiosis (interaction between organisms with net gain of fitness for both) and parasitism (reduced fitness of plant for benefit of microbe) (Takken & Rep, 2010). Plants and microorganisms have competing interests in the case of parasitism which leads to an evolutionary 'arms race' in which both sides are evolving to improve fitness in response to each other (Takken & Rep, 2010). These constant genetic changes lead to new resistant hosts and new virulent pathogen strains.

1.10.1 Plant immunity and resistance

The plant innate immune system is triggered into eliciting local and systemic defence responses when challenged by fungal plant pathogens (Selin *et al.*, 2016). This first line of defence is triggered by essential conserved and slow evolving molecules in many microbes, such as elongation factors and fungal cell wall components such as polysaccharides and chitin, which are collectively known as microbe or pathogen associated molecular patterns (MAMPs or PAMPs) (Newman *et al.*, 2013). They are recognised by transmembrane pattern recognition receptors (PRRs) on the surface of plant cells, resulting in MAMP or PAMP triggered immunity (the first phase of plant immunity), which can result in halting further pathogen colonisation (Jones & Dangl, 2006). MAMP (PAMP) induced defence responses include the production of reactive oxygen species, reactive nitrogen species (such as nitric oxide), alterations in the plant cell wall and synthesis of pathogenesis related (PR) proteins (e.g. enzymes, such as glucanase and chitinase, which directly attack pathogen structures) (Newman *et al.*, 2013).

To establish infection, fungal (and other) pathogens have to overcome this first line of defence, which they do through secretion of effector proteins. This has resulted in the plant defence system evolving a second line of defence called effector triggered immunity (ETI) (Figure 1.7) (Selin *et al.*, 2016). Effectors that elicit this response are recognised by plant resistance proteins (R proteins) which mostly encode for intracellular nucleotide-binding leucine rich repeat (NB-LRR) receptors. Effector triggered immunity is an accelerated and amplified PAMP triggered immunity response, resulting in disease resistance and localised cell death or hypersensitive response (HR) at the infection site (Jones & Dangl, 2006). This type of immunity is usually host specific and is often called a gene-for-gene interaction, where an effector protein (avirulence, Avr) is recognised by the corresponding R protein produced by the plant (Selin *et al.*, 2016).

1.10.2 *F. oxysporum* effectors

Effectors can be defined as "molecules that alter host cell structure and function, thereby facilitating infection (virulence factors or toxins) and/or triggering defence responses (avirulence factors or elicitors)" (Kamoun, 2009). Effectors can be distinguished into apoplastic effectors that are secreted into the plant extracellular space and cytoplasmic effectors that are translocated inside the plant cell (Kamoun, 2007). Pathogens and hosts are in an evolutionary 'arms race' in terms of effectors and R proteins, in which genetic changes that enhance fitness (either avoiding host detection or regaining pathogen recognition ability) are retained in the populations of both (Takken & Rep, 2010). This evolutionary adaptation of effectors and host resistance leads to new races of pathogens emerging along with subsequent resistant host cultivars (Takken & Rep, 2010). A well-documented example of this is the arms race between *Fusarium oxysporum* f. sp. *lycopersici* (FOL) and tomato.

In tomato, R genes that confer resistance against the wilt inducing FOL pathogen are called *I* (for *Immunity*) genes and have been introgressed into commercial cultivars. Several FOL races have been identified based on their ability to overcome *I* (and *I-1*), *I-2* and *I-3* resistance genes (Takken & Rep, 2010) and their corresponding avirulence genes identified. These are small secreted *in planta* effectors called Secreted in xylem (Six) proteins, coded for by *Secreted In Xylem* (*SIX*) genes (Rep *et al.*, 2004). Six4 (Avr1)

is recognised by *I* or *I-1*, Six3 (Avr2) is recognised by *I-2* and Six1 (Avr3) is recognised by *I-3* (Rep *et al.*, 2005). The presence/absence or mutation of these genes in FOL is thought to coincide with the evolution of new races. As *SIX4* is only present in FOL race 1 it is presumed that race 2 strains evolved from race 1 through loss of *SIX4*, while race 3 strains evolved from race 2 strains through point mutations in *SIX3* (Houterman *et al.*, 2009).

There are 14 characterised *SIX* genes in FOL (Houterman *et al.*, 2007; Lievens *et al.*, 2009; Schmidt *et al.*, 2013), identified through proteomics (mass spectrometry) of infected xylem sap and whole genome sequencing of isolates. Their role in pathogenicity has only been demonstrated for *SIX1*, *SIX3*, *SIX4*, *SIX5* and *SIX6* (de Sain & Rep, 2015), through gene knockout studies. All *SIX* genes (other than *SIX13*) are located on chromosome 14, a lineage specific (LS) chromosome found only in FOL, termed the pathogenicity chromosome due to its ability to confer virulence in a non-pathogenic *F. oxysporum* isolate when artificially transferred (Ma *et al.*, 2010; Schmidt *et al.*, 2013). The origin of the LS regions in FOL is thought to be from horizontal acquisition from another *Fusarium* species (Ma *et al.*, 2010), indicating that these regions can transfer pathogenicity related genes between species and may provide a mechanism for the evolution of pathogenicity in f. spp. The majority of the *SIX* genes in FOL were also identified in close proximity to specific transposable elements called miniature impalas (mimps), which allow the movement or deletion of genes in the genome (Schmidt *et al.*, 2013). Recombination by mimps might represent a mechanism of genetic variation in asexual fungi (Schmidt *et al.*, 2013) and could also be important in rearranging horizontally transferred genetic material to enable the evolution of new f. spp. and races. These LS regions are also high in mimps, transcription factors and other predicted effectors (Ma *et al.*, 2010) and searching for these genomic features in the genome can help uncover additional potential effectors (Schmidt *et al.*, 2013).

SIX genes have also been identified in other *F. oxysporum* f. spp. including f. sp. *cubense*, f. sp. *cepae*, f. sp. *niveum*, f. sp. *pisi*, f. sp. *melonis* and f. sp. *vasinfectum* (Fraser-Smith *et al.*, 2014; Meldrum *et al.*, 2012; Sasaki *et al.*, 2015; Taylor *et al.*, 2016; Covey *et al.*, 2014; Chakrabarti *et al.*, 2011). The presence/absence and sequence variation within *SIX* genes has been used as a novel approach to distinguish between *F. oxysporum* f. spp. and even between races. FOL is distinguished from other f. spp. as it is the only f. sp. to

contain *SIX1* - *SIX5* (Lievens *et al.*, 2009). Sequences of *SIX6* and *SIX7* were also found to vary between f. spp. but not within f. sp. (Lievens *et al.*, 2009). Similarly, sequence variation in *SIX8* in *F. oxysporum* f. sp. *cubense* (causing wilt in banana) has enabled race 4 to be distinguished from races 1 and 2 (Fraser-Smith *et al.*, 2014). Overall, *SIX* gene presence/absence and sequence variation could be used as a means of rapidly identifying *F. oxysporum* f. spp. and races.

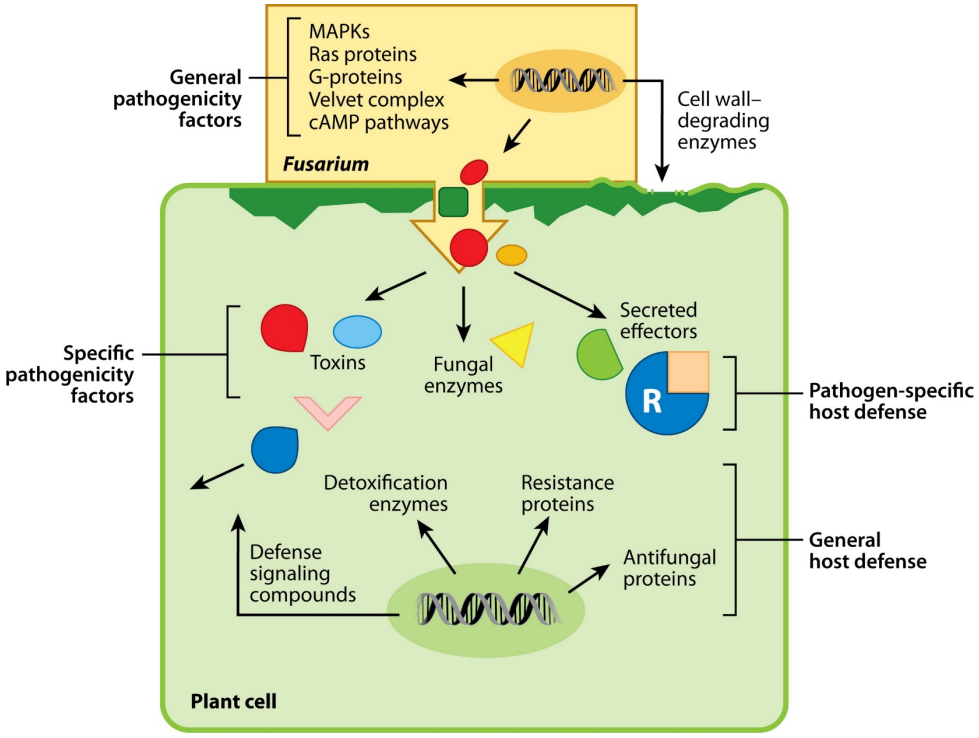


Figure 1.7 *Fusarium* pathogenicity and host defence responses including general cell wall degrading enzymes, specific secreted effectors and host resistance proteins (Ma *et al.*, 2013)

1.11 Genome analysis of *F. oxysporum*

The advent of next generation sequencing technologies (NGS) has accelerated the progress of genomic research as it is significantly cheaper and quicker than the previous Sanger technology (Giordano *et al.*, 2017), and can be used to determine structure, function and evolution of genomes. NGS sequencing technologies (such as Illumina) usually shear DNA into small fragments before being amplified and sequenced, and therefore sequence lengths tend to be quite short, between 50-400 bases (Giordano *et al.*,

2017). When these are assembled based on overlapping fragments, and multiple fragments cover the same regions, they tend to result in contigs which are much smaller than chromosomes; hence complex regions of the genome such as repetitive regions cannot be resolved if they are longer than the fragment length (Giordano *et al.*, 2017). To overcome this problem, long read sequencing technologies have recently been developed which provide read lengths up to hundreds of kilobases (kb), such as Pacific Biosciences and Oxford Nanopore (Weirather *et al.*, 2017). Although these long read technologies are useful for spanning repetitive regions and for de-novo genome assembly, they do have a much higher error rate than short read technologies such as Illumina (Weirather *et al.*, 2017). Genome sequencing and bioinformatics has been vital for progressing understanding of *F. oxysporum* interactions with their respective hosts by providing the data and tools for predicting effector genes and identifying resistance genes (Selin *et al.*, 2016). Genome sequencing of multiple f. spp. has been conducted in recent years, such as f. sp. *cepae* (Armitage *et al.*, 2018), legume hosts f. sp. *ciceris* and f. sp. *medicaginis* (Williams *et al.*, 2016), f. sp. *cubense* (Guo *et al.*, 2014), f. sp. *melonis* (Schmidt *et al.*, 2016) and the most widely studied f. sp. *lycopersici* (Ma *et al.*, 2010). Fungal effectors can be difficult to predict as they lack unifying sequence features within and across species, which can be attributed to rapid divergence and host specialisation (Sperschneider *et al.*, 2015). Therefore, effector prediction is based on broad criteria of a secretion signal, their size, and whether they are rich in cysteine residues (Sperschneider *et al.*, 2015). Understanding their location in the genome can also help effector prediction; for example, chromosome 14 in FOL is rich in effectors and can induce pathogenicity in non-pathogenic *F. oxysporum* isolates when transferred (Ma *et al.*, 2010).

Incorrect gene annotation can hamper effector prediction, but can be improved by mapping RNA sequencing (RNAseq) data to the reference genome, which helps to distinguish exon-intron boundaries and transcription start/stop sites (Gibriel *et al.*, 2016). Expression data is also enormously valuable in determining whether predicted effectors could actually play a role in pathogenicity. Transcriptomics has already been widely reported for *F. oxysporum* f. spp. supported by whole genome sequencing, with examples including f. sp. *melonis* (Sebastiani *et al.*, 2017), f. sp. *cubense* (Guo *et al.*, 2014), f. sp. *medicaginis* (Thatcher *et al.*, 2016) and f. sp. *cepae* (Armitage *et al.*, 2018). One of the best methods to validate gene models and expression data is through the use of proteomics, especially for those genes encoding secreted proteins and showing high

expression (Gibriel *et al.*, 2016). Expression of genes alone does not mean they form functional proteins, so the addition of protein analyses and proteomics is invaluable (Pandey & Mann, 2000). Some of the Six proteins in FOL have been characterised in the xylem sap of tomato plants, providing additional evidence they are involved in pathogenicity (Houterman *et al.*, 2007). Proteomics can also be used to investigate anti-pathogen activity in infected plants to discover immunity related proteins in the host, which could help to identify resistance mechanisms that can be bred into resistant cultivars (Castillejo *et al.*, 2015).

1.12 Project aims and objectives

The overall aim of this research was to determine the *Fusarium* species causing disease in UK peas and to understand the genetic basis for pathogenicity of different *F. oxysporum* f. sp. *pisi* (FOP) races.

The specific objectives were to:

1. Identify the different *Fusarium* species affecting peas in the UK and determine the importance of *F. oxysporum* as a pathogen.

2. Identify putative effector genes in three races of FOP using whole genome sequencing.

3. Evaluate the expression of putative effector genes in FOP races *in planta*

4. Verify the race type of pathogenic FOP isolates using pathogenicity tests with pea differential cultivars.

1.13 Thesis structure

Chapter 2 – Identification of *Fusarium* species affecting pea in the UK and importance of *F. oxysporum* as a pathogen

This chapter explores the different *Fusarium* species associated with diseased peas in UK fields and describes the use of molecular and phylogenetic analyses to classify these in relation to FOP isolates. Multiple pathogenicity assays were also assessed to determine the pathogenicity of *F. oxysporum* and FOP.

Chapter 3 – Prediction of putative effector genes in races of FOP using whole genome sequencing

Whole genome sequencing and assembly was conducted for three FOP isolates. Synteny between the FOP genomes and a reference *F. oxysporum* f. sp. *lycopersici* genome was explored and the presence and distribution of putative effector genes determined.

Chapter 4 – *SIX* gene profiling for determining FOP race type and effector expression *in planta*

SIX gene profiling was described as a potential molecular technique for distinguishing FOP races. In addition, qPCR was carried out to examine *SIX* gene expression *in planta* as well as RNAseq analyses to determine expression of genes following infection of pea by different FOP races.

Chapter 5 – Pathogenicity of *F. oxysporum* isolates in pea

This chapter describes the use of pathogenicity tests from Chapter 2 to distinguish between *F. oxysporum* isolates causing root rot and wilt symptoms in pea. FOP isolate race type was determined using pea differential cultivars.

Chapter 6 – General Discussion

This chapter presents overall conclusions and discussion of the main results from the thesis, outlining areas of improvement and future work.

2. Identification of *Fusarium* species affecting pea in the UK, and assessment of methods for determining pathogenicity of *F. oxysporum* on pea

2.1 Introduction

In investigations of pea root rot diseases worldwide, many *Fusarium* species have been associated with symptoms of root rot comprising black discolouration of the roots and stem base, yellowing of lower leaves, and stunted growth (Persson *et al.*, 1997). A range of *Fusarium* species are also commonly isolated from roots and the surrounding soil (Oyarzun *et al.*, 1993; Skovgaard *et al.*, 2002). For many years, *F. solani* was thought to be the major pea root rot pathogen (Chittem *et al.*, 2015), but several surveys have revealed that it is not always the most frequently isolated species. For instance, a study of root rot pathogens of pea in Canada found that *Fusarium* species were the most frequently isolated, with the highest proportion belonging to *F. oxysporum* (Hwang & Chang, 1989). However, subsequent studies in Canada identified *F. avenaceum* as the most prevalent species isolated from diseased pea roots (Fernandez, 2007). Similarly, findings from North Dakota in 2004/5 and 2008/9 also report *F. avenaceum* and *F. oxysporum* as the most frequently isolated species, along with other species such as *F. solani* and *F. redolens* (Chittem *et al.*, 2015). Research on root rot pathogens carried out in Denmark, Sweden and the Netherlands, found that *F. oxysporum*, *F. solani* and *F. avenaceum* were the most prevalent species, supporting results seen in other countries (Oyarzun *et al.*, 1993; Persson *et al.*, 1997). In the UK, *F. solani* and *F. oxysporum* were the most common species isolated from plants displaying root rot symptoms collected from 20 sites across Nottinghamshire, Lincolnshire and Worcestershire (Clarkson, 1978). This is similar to previous findings where *F. solani*, *F. oxysporum* and *F. redolens* were isolated from diseased peas across the major pea growing areas of the UK (Buxton, 1955). In most studies, pathogenicity tests were carried out with varying results, suggesting that not all species isolated are highly pathogenic (Chittem *et al.*, 2015; Persson *et al.*, 1997; Skovgaard *et al.*, 2002). Additionally, in most studies isolates were identified using morphological techniques only, reducing the accuracy and reliability of the results.

Fusarium oxysporum f. sp. *pisi* (FOP) was first described and distinguished from Fusarium root rot in 1925, by Jones and Linford, after it was found in around 50 fields in Wisconsin (Jones & Linford, 1925; Kraft, 1994). Symptoms differ from those caused by root rot pathogens in that leaves become dry and wilted, progressing from the lower leaves to the apex, until the whole plant was wilted and dried (Kraft & Pfleger, 2001). It was later assigned the name of FOP race 1 and has since been controlled using resistant varieties. Race 2 was described in 1933 as it was capable of causing disease in pea plants resistant to race 1, and was found to be as widespread as race 1 (Kraft, 1994). Race 5 was not identified until the late 1960's, where commercial cultivars resistant to race 1 and 2 were found to be susceptible in north-western Washington (Haglund & Kraft, 1970). The final race to be identified was race 6, found in western Washington, and was pathogenic on cultivars resistant to the other identified races (Haglund & Kraft, 1979). Race 6 was identified in 27% of the 640 fields sampled, with 40% containing race 5 isolates, and race 6 was distinguished from race 5 using two differential cultivars (Haglund & Kraft, 1979). The development of resistant varieties has enabled the successful control of all these four races of FOP infections on pea (Kraft, 1994). Early studies on FOP in the UK showed that some isolates of *F. oxysporum* collected from affected pea fields caused wilting similar to that observed previously for an isolate of race 1, while others showed root rot symptoms similar to those seen by *F. solani* (Buxton, 1955).

As with the pea root rot surveys, the majority of these studies with FOP have relied on morphological identification and race testing using a core set of pea differential cultivars (Kraft & Pfleger, 2001). Molecular techniques for identification are now common practice, using sequencing and phylogeny of housekeeping genes, which greatly improves the accuracy of species classification. In fungi, the internal transcribed spacer (*ITS*) is universally used for identification, but it has been shown to be inadequate for distinguishing different *Fusarium* species due to non-orthologous sequence types found within *Fusarium* species and within isolates (O'Donnell & Cigelnik, 1997; Summerell & Leslie, 2011). However, the *translation elongation factor 1*α (*TEF*) gene, which encodes an essential part of the protein translation machinery, resolves different *Fusarium* spp. and is therefore the marker of choice for identification. The gene has been shown to be consistently single copy in *Fusarium* and shows a high level of sequence polymorphism between closely related species (Geiser *et al.*, 2004). *TEF* sequencing was also used to investigate lineages within the *F. oxysporum* complex, but failed to reliably distinguish

between different *formae speciales* due to isolates being as genetically similar to other *formae speciales* as they are to isolates in their own *formae speciales* (O'Donnell *et al.*, 1998). Other housekeeping genes such as *β-tubulin* (*TUB2*) and *RNA polymerase II second largest subunit* (*RPB2*) have also been used to provide a more robust phylogeny within species (O'Donnell & Cigelnik, 1997; O'Donnell *et al.*, 2010).

Although sequencing of these genes is a useful tool for identification and understanding the phylogeny of the genus, they still do not allow different *F. oxysporum formae speciales*, races or pathogenic/non-pathogenic isolates to be distinguished. Molecular discrimination between *F. oxysporum* isolates is very difficult due to the polyphyletic nature of the many *formae speciales*, meaning that isolates from different *formae speciales* may be more closely related to each other than to isolates of the same *forma specialis* (Lievens *et al.*, 2009). More recently however, progress has been made through the use of effector gene presence / absence or sequence for distinguishing between certain *formae speciales*, between pathogenic and non-pathogenic isolates, and even between races of *F. oxysporum* f. sp. *lycopersici* (FOL) (Lievens *et al.*, 2009). However, in many cases, traditional pathogenicity tests are still required to confirm the *formae speciales*, race and pathogenicity of *F. oxysporum*. In FOP, there is currently no molecular way of distinguishing the four races, which still relies on pathogenicity tests using the pea differential cultivars.

There are a number of ways of assessing pathogenicity of root rot and wilt pathogens in pea. A widely used pathogenicity test is the root trim and dip method, used to distinguish pathogenic races of FOP (Haglund, 1989; Kraft, 1994). Although this assay has also been used to test the pathogenicity of *Fusarium* root rot isolates (Persson *et al.*, 1997), it is predominantly used for testing isolates causing Fusarium wilt. Pathogenicity tests with *Fusarium* isolates affecting pea have included tube tests with liquid inoculum (Dyer & Ingram, 1990), a sand-cornmeal inoculum layer between layers of vermiculite (Chittem *et al.*, 2015), sterilised soil inoculated with liquid inoculum (Clarkson, 1978) and soaking germinated seeds in a conidial suspension (Feng *et al.*, 2010). Each test appears successful in that the pathogenicity of different isolates can be assessed on the pea plants, allowing visible symptoms to be scored. Other tests used to evaluate the pathogenicity of other *F. oxysporum* f. spp., involving inoculating seeds and soil with spores or solid inoculum,

could also be applied to assess root rotting effects of *F. oxysporum* in pea (Taylor *et al.*, 2013).

The main aim of this chapter was to identify and characterise *Fusarium* species affecting peas in the UK and develop methods for investigating pathogenicity.

The specific objectives were:

1. To obtain *Fusarium* isolates from pea plants displaying symptoms of root rot/wilt from UK pea fields
2. To identify and characterise *Fusarium* species using common fungal housekeeping genes (*TEF*, *TUB2*, *RBP2*)
3. To assess methods for testing the pathogenicity of different isolates of *F. oxysporum* on pea

2.2 Materials and Methods

2.2.1 Sampling of pea fields and *Fusarium* isolate collection

Fusarium isolates were obtained from diseased pea plants, sampled from 74 fields across three of the major pea growing areas of the UK; Yorkshire, Lincolnshire and Suffolk. Plants collected for isolations showed symptoms of root rot or wilt near the end of the growing seasons (July-August) of 2015 and 2016. Sampling conducted in 2015 was more extensive, with the aim of collecting root rot and wilt causing *Fusarium* species by isolating from root sections of plants displaying root discolouration and stem wilt. However, in 2016 this was narrowed to mainly focus on wilt causing *Fusarium* species by isolating from mainly the stems of diseased plants displaying yellowing leaves/stems and wilt symptoms.

Fusarium cultures from sampled plants were obtained by surface sterilising root and lower stem sections (between stem base and 2nd - 3rd node) of five plants per field in a 5 – 10% bleach/sterile water solution (v/v) for 3 minutes. 5% bleach (4.5 g 100 mL^{-1} sodium hypochlorite, 1 – 5% available chlorine) in sterile water (v/v) was used to sterilise root sections from 2015 and 2016, and 10% bleach/sterile water (v/v) was used for 2016 stem samples after unsuccessful *Fusarium* isolation by sterilising with 5% bleach/sterile water. Root/stem sections were then washed twice in sterile water to remove bleach residue. Three sections from root/stem material were plated onto a full strength (39 g L^{-1}) potato dextrose agar plate (PDA; Merck, UK) containing 2 mL L^{-1} chlortetracycline (10 mg mL^{-1} in 1/1 methanol/water (v/v)) and incubated at 20˚C for around 5 days. Isolates which were distinct in morphology from the same field were sub-cultured onto PDA plates, and grown for approx. 7 days at 20˚C.

Additional isolates were also obtained in the form of agar plates and plant samples from The Processors and Growers Research Organisation (PGRO), derived from pea seed, plant and crop clinic samples from across the UK (collected 2012 – 2014). Further to this, isolates of known races of FOP were obtained from different countries, mainly the USA, Netherlands and Czech Republic. These were sub-cultured onto PDA and grown for 5-7 days at 20˚C.

2.2.1.1 Isolate selection and storage

All *Fusarium* isolates received from PGRO and overseas, and a selection of isolates obtained from the field sampling were taken forward for molecular identification. For the 2015 field samples, isolations were made from diseased root sections only, in order to capture any *Fusarium* isolate causing disease, and the morphology of cultures was recorded and grouped into categories based on colour and mycelial structure. Two isolates from the most commonly occurring categories per field were then chosen for identification, and if no common morphology was observed, two isolates were selected at random. For 2016 samples, isolations were made from both root and stem sections (2.2.1). However, only isolates obtained from stem sections were selected for molecular identification with the aim of specifically selecting those causing wilt rather than root rot.

All *Fusarium* isolates obtained were used to produce PDA slope cultures for storage at 4˚C and spore suspensions in potato dextrose broth (PDB) + 20% glycerol (v/v) for storage on ceramic beads at -80˚C.

2.2.2 Molecular identification of *Fusarium* isolates

All selected *Fusarium* isolates were identified and characterised through *TEF* sequencing. Three 5 mm agar plugs were removed from the growing edge of each actively growing culture and used to inoculate 20 mL of 50% potato dextrose broth (PDB) in a 50 mL tube or in a Petri dish. Tubes/plates were incubated at 20˚C for 5 days, with the tubes positioned at a 45˚ angle. PDB was then removed by centrifugation (3000 rpm for 15 minutes) and mycelium from each isolate was rinsed twice with sterile water (centrifugation at 3000 rpm for 15 min each time). The remaining mycelium was freeze-dried for 2 days.

DNA was extracted from freeze-dried mycelium either using a DNeasy plant mini kit (Qiagen, Hilden, Germany) or a rapid DNA extraction protocol (Acme) (S. Rehner, personal communication). The DNeasy plant mini kit was used in accordance with manufacturer's protocol with a minor modification whereby the mycelium was homogenised in a lysing matrix-A tube (MP Biomedicals, CA, USA) in a FastPrep-24^{TM} machine set at 6 ms^{-1} for 40 s. For the Acme protocol, the mycelium was transferred into

2 mL tubes containing 6-10 glass beads and silica beads (0.1mm, BioSpec Products) and ground three times in a FastPrep-24™ machine set at 5.5 ms^{-1} for 20 s. Acme DNA extraction reagent (300µl of sodium metasilicate 2.1g, citric acid 0.5g, 2-butoxy ethanol 2.64 mL, 1M Tris-HCl pH 7.0 13.6 mL) was added, tubes heated for 10 minutes at 100˚C, and centrifuged at 16,000 g for 5 minutes, rotated 180˚ and centrifuged again for 5 minutes. The supernatant (175 µl) was transferred to a clean tube and diluted 1:10 in TE buffer for use in PCR

Identification of *Fusarium* isolates was carried out by PCR amplification of the *TEF* gene using published primers (Vágány, 2012) (Table A 4.1). All PCR reactions were set up using REDTaq® ReadyMix® (Sigma-Aldrich) in 20 µL reaction volumes containing 1 µL of DNA and a final concentration of 0.5 µM of each primer. Thermocycling conditions for *TEF* were: one cycle of 2 min at 94˚C; 30 cycles of 45 s at 94˚C, 30 s at 64˚C and 1 min at 72˚C, followed by one cycle of 5 min at 72˚C. PCR amplicons were visualised using gel electrophoresis (1% agarose gel containing GelRed™ at 2 µL per 100 mL of gel), purified using the QIAquick PCR Purification Kit (Qiagen) and sequenced (Eurofins, Germany, via the University of Warwick Genomics Facility) using the forward primer sequence. Sequences were subjected to BLAST searches (Basic Local Alignment Search Tool) using the National Centre of Biotechnology Information (NCBI) database to identify species based on sequence identity values.

TEF sequences from all identified *Fusarium* species from diseased peas in UK fields were aligned and trimmed in MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets (Kumar *et al.*, 2016). 110 isolates were selected and used to construct a maximum-likelihood tree using a Kimura 2-parameter plus gamma model (Kimura, 1980).

Isolates of *F. oxysporum* only were used to construct a further maximum likelihood tree using *TEF* sequences from 81 *F. oxysporum* isolates from diseased peas in UK fields, 27 *F. oxysporum* isolates from other *formae speciales* obtained from the Broad Institute by Taylor *et al.* (2016) and 32 previously FOP race typed isolates from overseas. As before, sequences were aligned and trimmed in MEGA 7.0 and then used to construct the tree using a Kimura 2-parameter model (Kimura, 1980).

Based on the phylogeny of *F. oxysporum* isolates using *TEF* above, a selection of the

isolates obtained from diseased peas in UK fields and the FOP race typed isolates from overseas were further characterized through sequencing of two additional housekeeping genes, *RPB2* and *TUB2*. Both sets of primers (Table A 4.1) used the same thermocycling conditions as *TEF* but with an annealing temperature of 60˚C. Products were visualised on a gel and purified before sequencing, as above. Sequences obtained were used to construct a concatenated maximum-likelihood tree using the calculated best model, Kimura 2-parameter plus gamma model (Kimura, 1980).

Bootstrap values were inferred from 1000 replicates (Felsenstein, 1985) for all three trees produced and displayed next to the relevant branch.

2.2.3 Assessment of inoculation methods to determine pathogenicity of *F. oxysporum* isolates on pea

A range of methods were tested to determine whether the pathogenicity of *F. oxysporum* isolates differed according to the type of inoculation. The inoculation methods included soaking seeds in spore solutions, inoculating pea seedlings grown in agar in test tubes with liquid inoculum, and finally trimming the roots of seedlings before dipping in spore suspensions.

2.2.3.1 Seed inoculation pathogenicity test

F. oxysporum isolates (five) were assessed for their pathogenicity in a seed inoculation test adapted from Taylor *et al.* (2013). Cultures of PG18, PG19, PG4, PG3, PG21 and Fo47 were initiated from glycerol bead stocks onto PDA and grown for 14 days at 20˚C. Spores were removed from the plates by pouring 10 mL of sterile distilled water onto one of four plates, disturbing all the spores and mycelium using a sterile plastic spreader and then pipetting the liquid onto the next plate. The solution was sieved through three layers of muslin (autoclaved) to filter out the mycelium and the concentration of the remaining spore suspension adjusted to 1 x 10^6 spores mL^{-1} in sterile water, using a haemocytometer.

Pea seeds (cv. Avola, Kings seeds, Colchester) were soaked in 10 mL of spore suspension for 2 hours before sowing into Levingtons F2 + S compost in 24 cell modular trays, one

tray per isolate. Trays were placed in a controlled temperature glasshouse (25˚C day, 18˚C night, 16 h photoperiod). There were four replicate trays per isolate and one tray of a non-inoculated control treatment (pea seeds soaked in sterile water only) was also included in each replicate. Trays were watered from below to prevent cross-contamination. Emergence and plant mortality were assessed twice a week for 28 days, or until the control plants began to senesce.

2.2.3.2 Test tube pathogenicity assay

F. oxysporum isolates (five) were assessed for their pathogenicity in an *in vitro* test tube assay originally implemented at PGRO. Agar (20 mL, 9.75 g agar powder (Fisher, UK) in 1 L tap water) was dispensed into large test tubes (25 mm x 150 mm), sealed with a non-absorbent cotton wool bung and covered in aluminium foil. The tubes were subsequently autoclaved (121˚C for 20 min) and cooled at a 4 cm slant (from horizontal). Pea seeds (cv. Avola) were surfaced sterilised in a 10% bleach (4.5 g 100 mL^{-1} sodium hypochlorite, 1 – 5% available chlorine) and sterile water solution (v/v) for 5 min and rinsed thoroughly with sterile distilled water until no bubbles remained. Under sterile conditions, one seed was then placed in each tube before it was resealed and incubated at 20˚C with a 16 h photoperiod for six days. Tubes were arranged in a randomised block design in the incubator. Spore suspensions of each *F. oxysporum* isolate (PG18, PG19, PG4, PG3, PG21and Fo47) were prepared as in 2.2.3.1 but in 0.1% tap water agar, adjusted to 5 x 10^5 spores mL^{-1} and 1 mL used to inoculate each tube. Non-inoculated control tubes received 1 mL of 0.1% tap water agar only. Pea seedlings were assessed for disease symptoms every 3-4 days post inoculation, using a scoring system developed at PGRO which assigns a root browning severity score between 0 and 6 (Figure 2.1).

Score	0	1	2	3	4	5	6
% root browning	0	<10	10-25	25-50	50-90	100	100
Description	Healthy	Slight	Slight-moderate	Moderate	Moderate-severe	Severe	Plant dead

Figure 2.1 Disease scoring system for test tube pathogenicity assay. Scores represent severity of root browning from 0 (no browning/healthy plant) to 5 (severe browning and wilting plant). The final score of 6 represents plant death.

2.2.3.3 Root dip pathogenicity test

The ability of *F. oxysporum* isolates to cause systemic wilt symptoms in pea when grown in vermiculite, compost and a 50/50 (v/v) mix of the two was assessed using a root dip test adapted from Bani *et al.* (2012). Pea seeds (cv. Avola) were sown in 24 cell modular trays containing vermiculite (1 – 3 mm diameter) until 2-3 nodes tall (approx. 14 days), in the glasshouse (25˚C day, 18˚C night, 16 h photoperiod). Roots were removed from vermiculite and trimmed by a third in length, before being immersed for 5 minutes in a spore suspension of each isolate. Spore suspensions were prepared as before by removing spores from *F. oxysporum* cultures (FOP1, FOP2, PG16, PG18 and Fo47) grown on PDA plates for 2 weeks at 20˚C using sterile distilled water and adjusted to 1 x 10^6 spores mL^{-1} of water. Control plant roots were immersed in sterile distilled water only. Following inoculation, plants were re-sown in individual pots (7 x 7 x 8 cm) containing vermiculite, compost (Levingtons F2 + S) or a 50/50 (v/v) mix of compost/vermiculite. Plants were maintained in the glasshouse (as above) and watered with the following: the vermiculite pots were supplemented with fertiliser (Vitax 2:1:4, diluted 100-fold) when watered every day, the 50/50 pots were watered every day and supplemented with fertiliser every other day, and the compost pots received only water every day. Disease symptoms were

assessed 1-2 times per week for 4-6 weeks depending on pea varieties, commencing 7 days after inoculation, by recording the proportion of leaves showing symptoms of wilt for each plant.

2.2.4 Statistical analyses

All statistical analyses were carried out in Genstat® (release 18.1, VSN international Ltd), with advice and support from Andrew Mead, Rothamsted Research. For the seed inoculation pathogenicity test, the percentage seed germination was transformed using a logit transformation due to the heterogeneity of variance in percentage data, before implementing an analysis of variance (ANOVA), considering the blocking structure of a randomised block design. Significant ANOVA results were interpreted by comparing ANOVA treatment means using the least significance difference values (LSD) at the 5% level. The table of means was back transformed to produce values on the percentage scale for plotting on a graph. The percentage survival data for the seed inoculation test was also analysed using a Generalised Linear Model (GLM) with logistic regression, with fitted terms of replicate (block) + treatment (isolate) and a dispersion parameter fixed at 1. Interpretations from the GLM were made by comparing t probabilities calculated with reference to the untreated control for each isolate.

For the test tube pathogenicity assay, the disease severity scores at each time point were analysed using ANOVA, considering the blocking structure of a randomised block design. Mean scores were compared using the LSD value (5%).

For the root dip pathogenicity test, the number of wilted leaves was calculated as a proportion of the total number and transformed using a logit transformation, to reduce the variation effects of percentage data, with the addition of an offset value of 1 to allow for any values of 0% or 100%. The transformed data was then analysed using a two-way ANOVA, comparing the different growth mediums with the proportion wilt for each isolate. Means for each isolate were compared to the control using the LSD at the 5% level, and were then back transformed before plotting on a graph.

2.3 Results

2.3.1 Sampling of pea fields and *Fusarium* isolate collection

During the growing seasons of 2015 and 2016 (fields sampled in July and August), a total of 504 *Fusarium* isolates were collected from root and stem sections of plants from 74 fields across Yorkshire, Lincolnshire and Suffolk (Figure 2.2 a). From these, 212 were selected for molecular identification using *TEF* gene sequencing as well as 88 isolates obtained from PGRO (Table A 2.2, Table A 2.1).

As outlined in Section 2.2.1.1, *Fusarium* isolates derived from field samples were selected for identification based on culture morphology. From this, isolates were assigned to 15 categories. Agar colour between categories ranged from pale yellow, through orange and red colours to deep purple. Aerial mycelium structure was classified based on whether it was "fluffy" or flat on the plate, spindly or displayed concentric circles. In general, most categories contained isolates of the same species (Table A 2.4), but the high colour variation and mycelial growth within each category made it difficult to identify *Fusarium* species without molecular identification.

2.3.2 Molecular identification of *Fusarium* isolates

Fusarium isolates were identified following DNA extraction, PCR and sequencing of the *TEF* gene. Of the 212 isolates selected from the field samples, 202 were successfully identified, along with all 88 of the additional samples from PGRO. The remaining 10 isolates could not be identified after repeated DNA extractions and PCR reactions with *TEF* and *ITS*. Of the 157 isolates derived from pea roots in 2015, the majority were identified as *F. oxysporum* (87 isolates) followed by 31 isolates for *F. solani* (Figure 2.2b). However, of the 45 isolates derived and identified from pea stems in 2016 the largest proportion of isolates were identified as *F. solani*, (30 isolates) followed by only four for *F. oxysporum* (Figure 2.2b). The majority of isolates from PGRO were identified as *F. redolens* (30 isolates) closely followed by *F. oxysporum* (29 isolates, Figure 2.2 b). Across all samples, five main species were identified: *F. oxysporum*, *F. solani*, *F. redolens*, *F. culmorum* and *F. avenaceum*. Other species identified (grouped into 'other *Fusarium* species' in Figure 2.2b) included *F. poae*, *F. equiseti*, *F. graminearum* and *F.*

lacertarum. Although all samples were collected from fields with peas showing similar disease symptoms (Figure 2.2a), there was a large variety of *Fusarium* species identified from all areas. Therefore, no correlation between plant symptoms and causative disease agent could be made. Overall, these results show that the largest proportion of isolates (40%) from all the samples collected were identified as *F. oxysporum*.

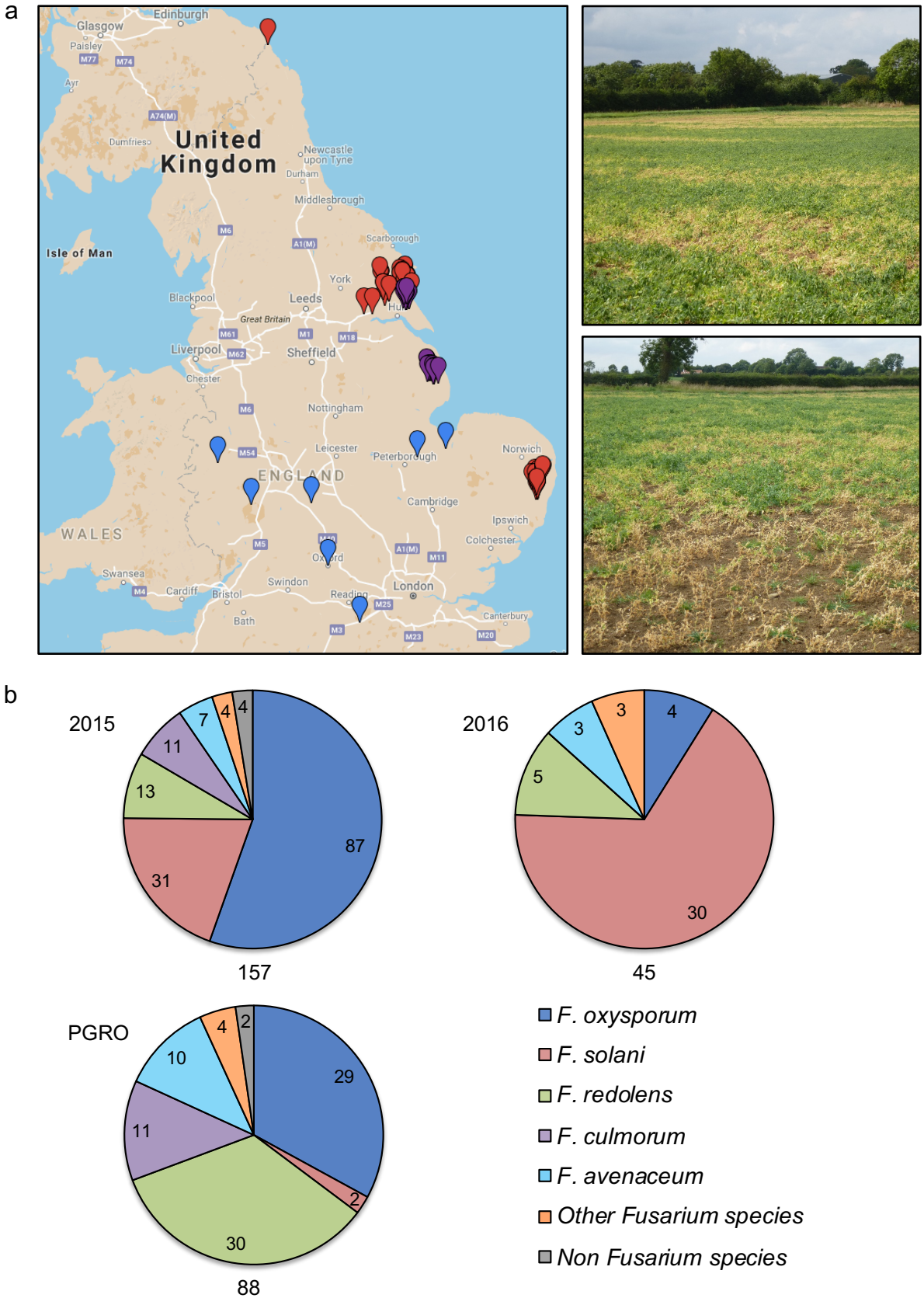


Figure 2.2 Sampling location and identification of isolates from UK fields. (A) Field locations of sampling sites (red pins = 2015 fields, purple pins = 2016 fields, blue pins = PGRO sampling sites/crop clinic locations) and symptoms in selected fields from Yorkshire. (B) Frequency of *Fusarium* species identified from UK fields in 2015, 2016 or from PGRO samples.

The *TEF* gene was also sequenced for isolates previously identified as FOP sourced from other countries, including the USA, Algeria and Czech Republic. Even though these isolates had supposedly been tested on pea differentials and had been assigned a FOP race type, only two of the four isolates from Algeria and only one of the three isolates from the Czech Republic were identified as *F. oxysporum* and were instead identified as *F. avenaceum* and *F. redolens*. In addition, one isolate from the USA was identified as *F. solani*. (Table 2.1). The three historic FOP isolates from Warwick HRI (FOP1, 2 and 5) were previously recorded as FOP races 1, 2 and 5 respectively (Linfield, 1994), and were used as the reference sequences for these three races.

Table 2.1 The location, source and number of isolates of each race obtained from the UK and abroad, used for molecular identification and phylogenetic analysis. The columns of number of isolates of each race refer to the race-type they had been assigned before they arrived. The column of isolates identified as *Fusarium oxysporum* (*F. oxysporum*) are the results of molecular identification after arrival.

No. isolates received	Original location	Source	No. of isolates of FOP races			No. identified as *F. oxysporum*
			Race 1	Race 2	Race 5/6	
4	Algeria	Dr A. Merzoug	2	1	1	2
3	CBS, The Netherlands	CBS collection	2	1		3
3	Czech Republic (via PGRO)	Dr L. Herold	2	1		1
2	USA	Dr R. J McGee	1	1		2
5	Athena, OR, USA	Dr L. Porter		5		5
6	Minnesota, USA	Dr L. Porter		6		6
2	MT. Vernon, USA	Dr L. Porter		2		2
1	Palouse, WA, USA	Dr L. Porter				0
3	Prosser, WA, USA	Dr L. Porter		3		3
3	Quincy, WA, USA	Dr L. Porter				3
1	Steptoe, USA	Dr L. Porter		1		1
4	Unknown	C. Linfield/ NIAB-EMR	2	1	1	4

All *TEF* sequences for the 202 isolates of different *Fusarium* species collected from diseased peas in UK fields and the 88 isolates of different *Fusarium* species from PGRO were aligned in MEGA 7.0 and a selection (110) used to construct a maximum likelihood tree. The tree shows clear separation of the different species into distinct clades, confirming the accuracy of the *TEF* sequencing (Figure 2.3). The tree was rooted using *Cylindrocarpon permirum* (GenBank accession: KJ022398.1) as both this genus and the *Fusarium* genus reside within the *Nectriaceae* family. There was clearly a diverse range of *Fusarium* species collected from diseased peas in UK fields. The *F. oxysporum* and *F. avenaceum* clades appear to contain isolates from multiple evolutional lineages, whereas the other clades contain no variation based on *TEF*.

A selection of *F. oxysporum* isolates (81) only from diseased peas in the UK and from PGRO were used to construct a separate maximum likelihood tree along with the FOP races from Algeria, Czech Republic, UK and USA and other *F. oxysporum formae speciales* (obtained from NCBI and Taylor *et al*., 2016). *F. oxysporum* isolates were separated into eight main clades (Figure 2.4), with the majority of the isolates collected from UK fields being grouped into clades 1, 2 and 8. These were distinct from the known races of FOP which formed separate clades, with most of the race 1 isolates being grouped into clade 6 and race 2 isolates mostly found in clade 7. Some of the FOP race 2 isolates grouped with the historic race 5 isolate (FOP5 isolate from Warwick HRI) in clade 3. The distinction of the isolates from diseased UK peas and the FOP race typed isolates mainly from the USA led us to hypothesise that the *F. oxysporum* isolates from UK fields were most likely not FOP isolates capable of causing wilt, and therefore were designated as causing root rot. Clade 4 contained a collection of other sequences obtained from NCBI and the Broad Institute by Dr Andrew Taylor, Warwick Crop Centre (Taylor *et al.*, 2016), and mostly comprised isolates of other *F. oxysporum formae speciales* such as f. sp. *cepae* (onion), *narcissi* (daffodil), *ciseris* (chickpea), *lycopersici* (tomato), *phaseoli* (bean), as well as some FOP isolates.

Figure 2.3 Maximum likelihood tree for selected *Fusarium* species collected from UK pea fields based on an alignment of *translation elongation factor 1α* (*TEF*). Numbers represent bootstrap values from 1000 replicates. Scale bar indicates 0.05 nucleotide substitutions per site. The tree is rooted through *Cylindrocarpon permirum* (GenBank accession: KJ022398.1).

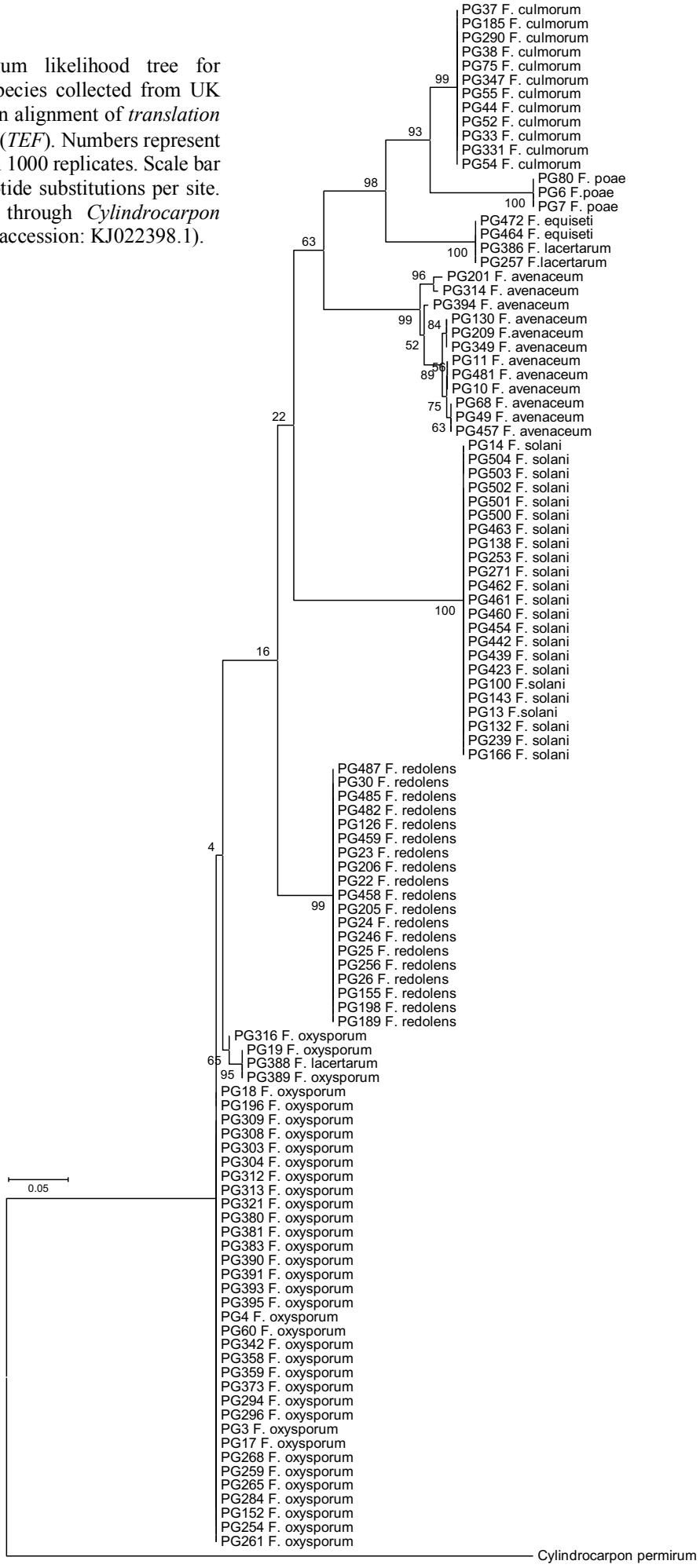
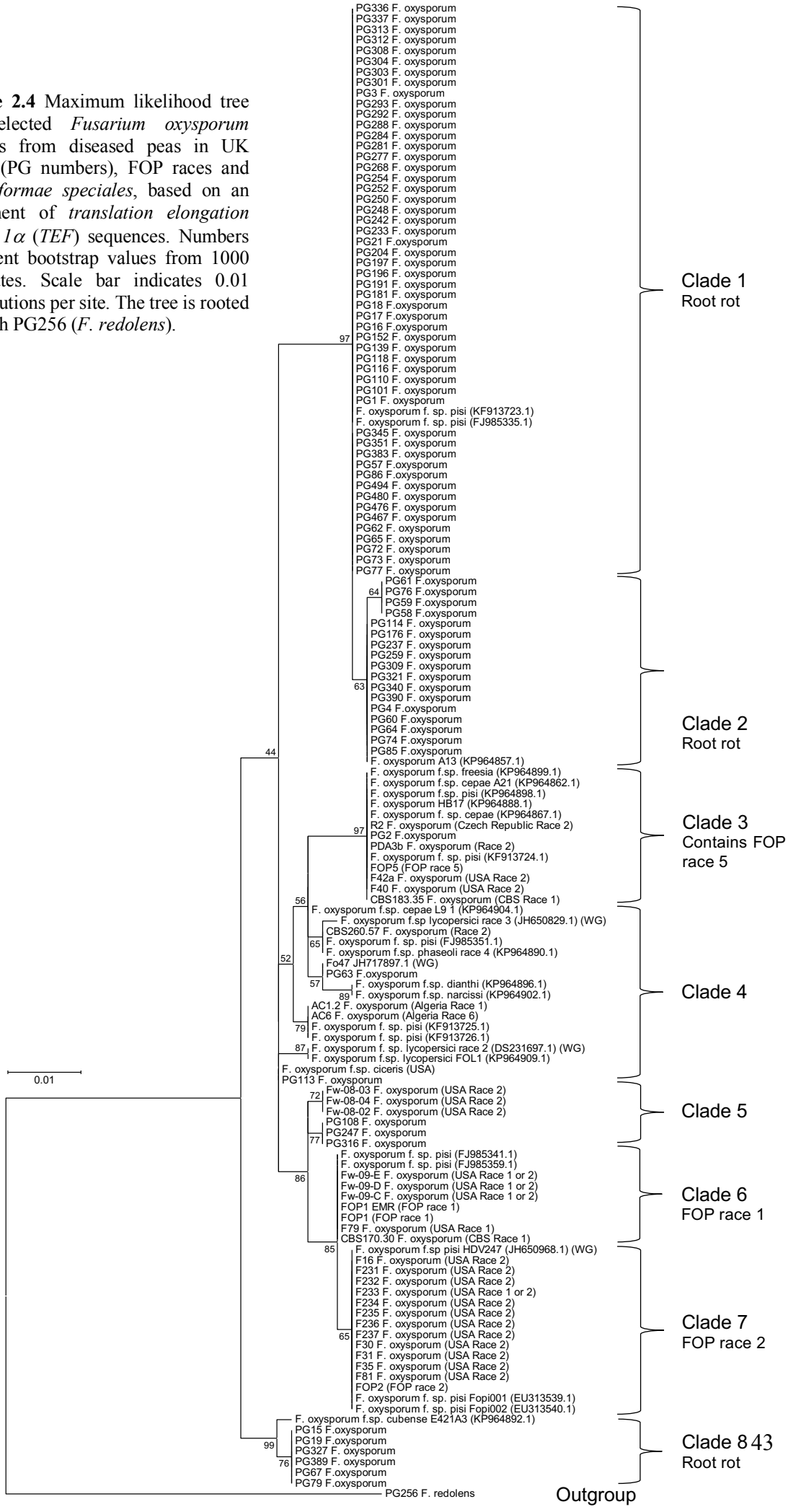


Figure 2.4 Maximum likelihood tree for selected *Fusarium oxysporum* isolates from diseased peas in UK fields (PG numbers), FOP races and other *formae speciales*, based on an alignment of *translation elongation factor 1α* (*TEF*) sequences. Numbers represent bootstrap values from 1000 replicates. Scale bar indicates 0.01 substitutions per site. The tree is rooted through PG256 (*F. redolens*).

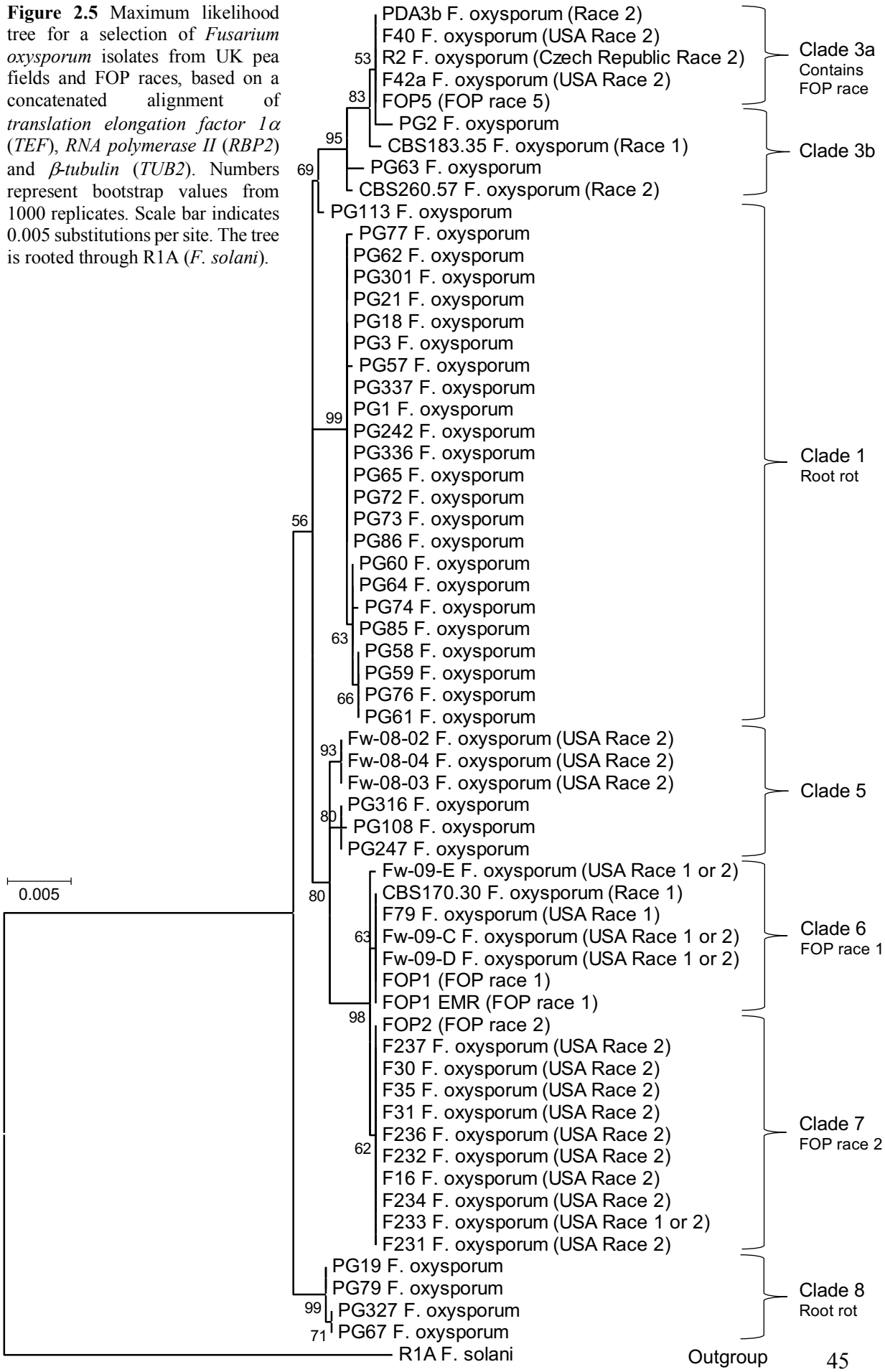


Two additional housekeeping genes, *RPB2* and *TUB2* were sequenced and used with *TEF* sequences to construct a concatenated maximum-likelihood tree for selected isolates of *F. oxysporum* (33 isolates from diseased UK peas and 28 previously race typed isolates). This tree reflected the results obtained using the *TEF* sequences alone (Figure 2.4) and grouped isolates into distinct clades. *F. oxysporum* isolates collected from diseased peas in UK fields were again separated into a different clade (clade 1) from the other FOP isolates (Figure 2.5). FOP race 1 and 2 isolates also grouped into separate clades (6 and 7 respectively), but both originated from the same node. Clade 3 from the *TEF* tree (Figure 2.4) was split into two clades as the additional concatenated gene sequences resulted in ingroup distinctions, which separated putative race 5 wilt isolates (clade 3a) from the other root rot causing isolates (clade 3b). Other smaller clades such as clades 5 and 8 contained the same isolates as those in respective clades of the *TEF* tree, e.g. PG247 and PG108 (clade 5) and PG19 and PG79 (clade 8) (Figure 2.5).

The similarities in clades between *TEF* and concatenated *F. oxysporum* phylogenies were supported by high bootstrap values with all (but one branch) being above 50%, confirming the reliability and accuracy of the positions of isolates in clades.

All trees were also constructed using the neighbour-joining and minimum-evolution methods and a similar topography was observed in all cases (data not shown).

Figure 2.5 Maximum likelihood tree for a selection of *Fusarium oxysporum* isolates from UK pea fields and FOP races, based on a concatenated alignment of *translation elongation factor 1α* (*TEF*), *RNA polymerase II* (*RBP2*) and *β-tubulin* (*TUB2*). Numbers represent bootstrap values from 1000 replicates. Scale bar indicates 0.005 substitutions per site. The tree is rooted through R1A (*F. solani*).



2.3.3 Assessment of inoculation methods to determine pathogenicity of *F. oxysporum* isolates on pea

Initially three plant/seed inoculation methods were tested on a small number of *Fusarium* isolates to determine which resulted in disease development. Firstly, a seed inoculation test, where seeds were soaked in spore solutions, secondly, an assay where seedlings were grown in agar filled test tubes and then inoculated with spore solutions, and finally, a root dip method where seedling roots were trimmed and immersed in spore solutions. At the time of testing, only a small number of isolates had been characterised (those obtained from PGRO). Isolate selection was based on previous work by an MSc student where pathogenicity of isolates were tested using two of the above inoculation methods (seed and agar inoculation tests) (Adedamola, 2014) and by using phylogenetic analysis (Figure 2.4).

2.3.3.1 Seed inoculation pathogenicity test

In the seed inoculation test, percentage germination was calculated to determine the level of pre-emergence damping off. ANOVA revealed that there were significant differences between isolate means, including the control ($p < 0.001$). Comparing logit transformed means with a 5% LSD of 0.71 (d.f.17) there were significant differences in the percentage germination for all *F. oxysporum* isolates compared to the control at 23 dpi. Percentage germination was significantly decreased for three isolates (PG21, PG3 and PG4) compared to the supposed non-pathogenic isolate Fo47, with two isolates PG18 and PG19 showing no significant difference (Figure 2.6a). Comparing back transformed means showed that the control treatment had the highest percentage germination at 74%, with PG19 the second highest at 49%, and PG21 (8%) causing the lowest percentage germination (Table 2.2). Results for percentage germination only changed marginally overtime from the first time point (7 dpi, data not shown), suggesting that seeds either germinate early on in the experiment or not at all.

The percentage survival of pea plants in the seed inoculation test were analysed using a General Linear Model (GLM) with logistic regression. Four *F. oxysporum* isolates (PG18, PG21, PG3 and PG4) resulted in significantly lower plant survival than the control

($p < 0.01$) at 23 dpi (Figure 2.6b). Comparing predicted means from the regression model for percentage survival, the control and Fo47 treatments had 97% and 98% survival respectively, compared to PG18 at 25% and PG21 at 43%. (Table 2.2). The percentage survival for each treatment was observed over the course of the experiment (Figure 2.6c), where it was observed that isolates PG4, PG18, PG19 and PG3 resulted in considerably decreased percentage plant survival after 17 dpi. After this time point the percentage survival inoculated with these four isolates did not decrease much further.

These preliminary results show that all isolates had a significant impact on the percentage of seeds that germinated, while some isolates also significantly reduced the survival of plants over time post-germination. Therefore, there are potentially two measures of pathogenicity for this test. Symptoms of root rot were clearly visible in plants inoculated with pathogenic isolates (PG18) compared to control plants (Figure 2.6d).

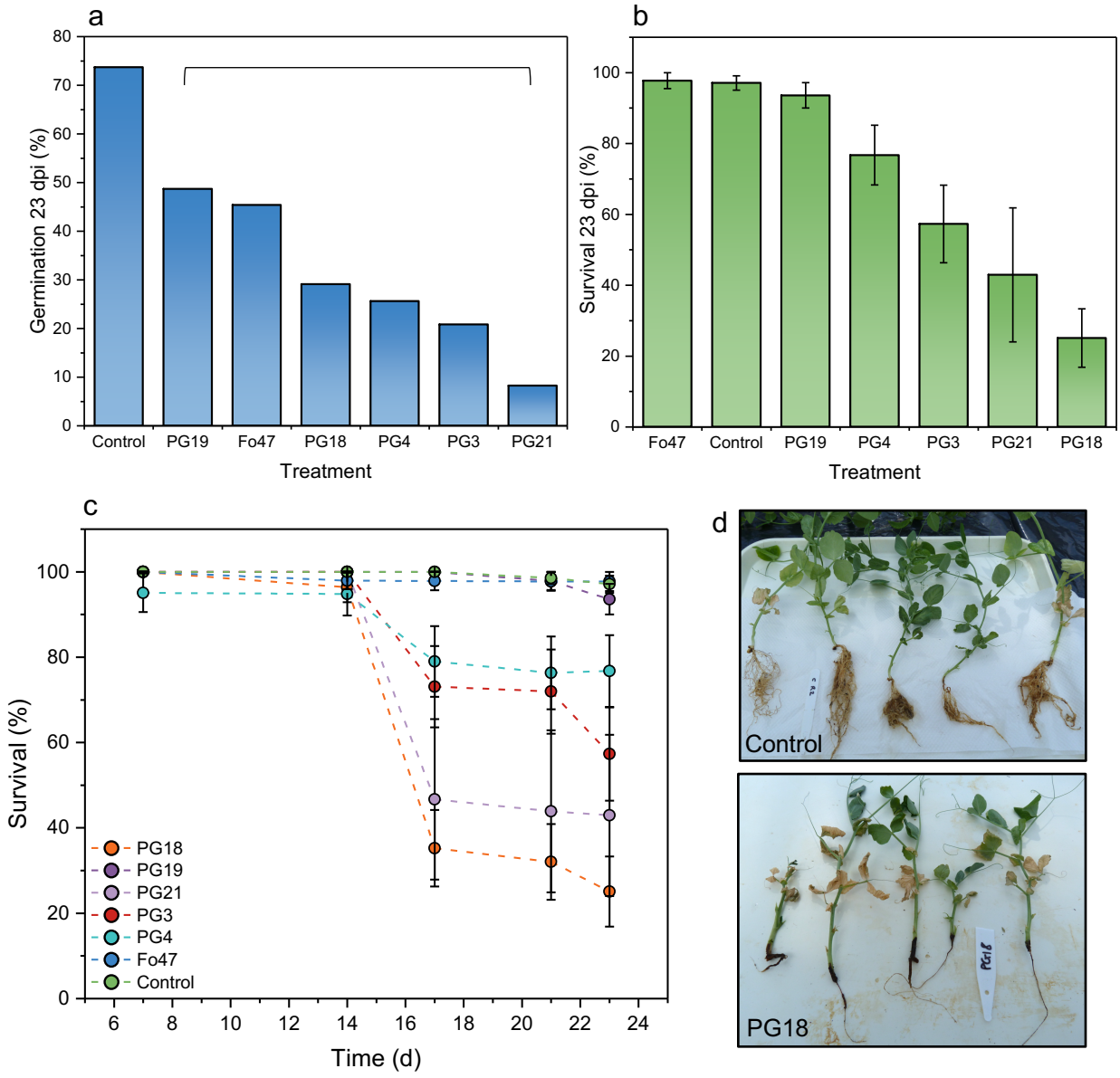


Figure 2.6 Preliminary results of the pathogenicity of six *Fusarium oxysporum* isolates using the seed inoculation method. a) back transformed means of percentage germination out of a total of 24 seeds sown at 23 dpi, all isolates (covered by the bracket) caused significantly less germination compared to the control; b) survival (back transformed predicted means) of seedlings as a percentage of those that germinated at 23 dpi; c) survival of seeds (back transformed predicted means) as a percentage of those that germinated over the time course of the experiment; d) symptoms of root rot in plants inoculated with PG18 compared to control plants, 23 dpi. Error bars for b) and c) represent predicted standard errors (back transformed from the table) for the percentage survival at 23 dpi.

Table 2.2 Summary of means used for statistical comparisons from transformed data for % germination of pea seeds inoculated with *Fusarium oxysporum* isolates 23 dpi, and estimates of parameters (from GLM) and their standard errors for % pea seedling survival at 23 dpi, using the seed inoculation test.

Isolate	ANOVA (germination)	GLM with logistic regression (survival)	
	%	% survival 23 dpi	Standard error
Control	1.032		
Fo47	-0.185	0.250	1.24
PG18	-0.892	-4.607	0.84
PG19	-0.052	-0.831	0.93
PG21	-2.409	-3.800	1.06
PG3	-1.337	-3.218	0.85
PG4	-1.067	-2.318	0.86
d.f	17		
5% LSD	0.7095		

2.3.3.2 Test tube pathogenicity assay

In the test tube pathogenicity assay, significant differences were observed between the inoculated plants and the non-inoculated plants ($p < 0.001$), with the control treatment causing no visible symptoms on the roots. All inoculated treatments however caused similar symptoms, and no isolate resulted in greater virulence (Figure 2.7). No significant differences occurred between inoculated treatments (means compared with 5% LSD), apart from Fo47 which caused significantly fewer symptoms on the pea roots compared to all other isolates. However, Fo47 still caused significantly more browning of the roots compared to the control. All isolates scored between 4.75 and 5.38 on average, whereas Fo47 scored only 1.62 on average at 14 days post inoculation. Scores for all isolates increased over time compared to the control (Figure 2.7), but this test did not distinguish between isolate pathogenicity at any time point.

Overall, this test therefore did not result in differences between pathogenic isolates but did allow potentially pathogenic isolates to be identified compared to the control.

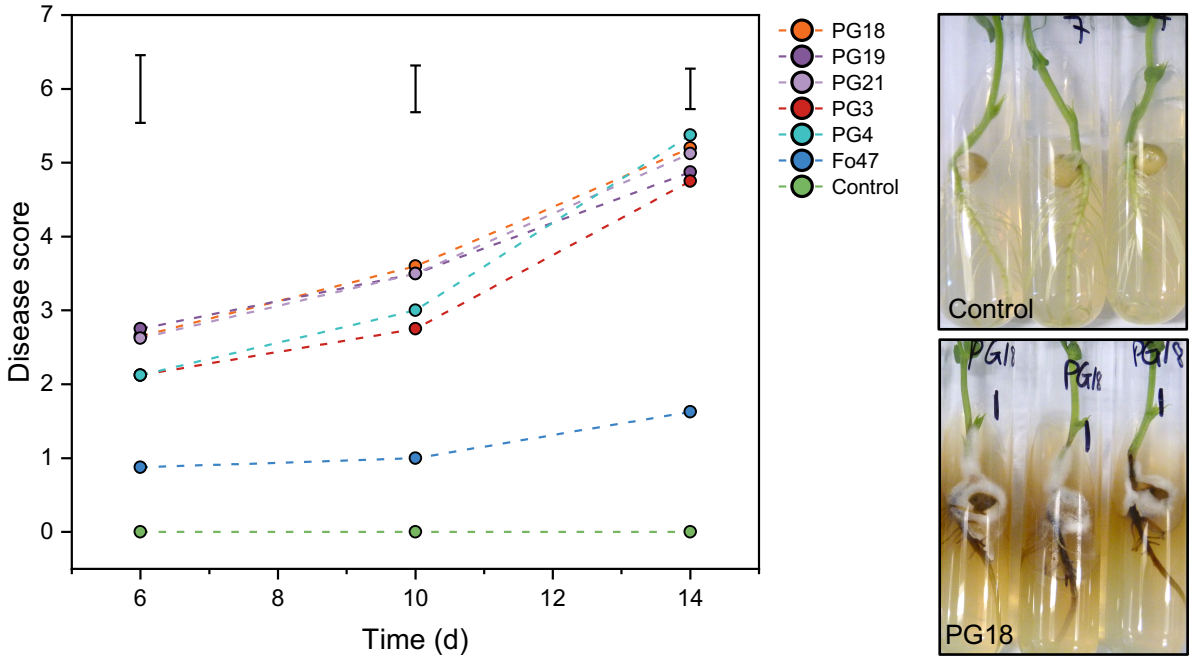


Figure 2.7 Average disease score of six *Fusarium oxysporum* isolates on pea seedlings grown in agar in test tubes, as part of the tube inoculation test. Disease symptoms were scored at 6, 10 and 14 dpi. Error bars represent the least significant difference (LSD) at 5% level for each time point. Pictures show root rot symptoms on plants inoculated with a pathogenic isolate (PG18) compared to the control.

2.3.3.3 Root dip pathogenicity test

In the root dip test designed to detect vascular wilt, two isolates of FOP (FOP1 and FOP2) and two isolates from diseased UK peas were used to test the differences between types of growth medium (vermiculite, compost and a 50/50 mix of both). Overall, there was a significant interaction between isolate and soil type (growth medium) for the proportion of leaves wilted ($p < 0.001$). FOP2 resulted in 100% of leaves being wilted by 29 dpi for all three soil types (Figure 2.8), and was the only isolate which resulted in significant wilt compared to the control (using 5% LSD, Table 2.3). PG18, PG16, FOP1 and Fo47 resulted in no significant increase of percentage wilt compared to the control for each soil type (compared using 5% LSD), with Fo47 and FOP1 resulting in significantly less wilt than the control when grown in compost (Table 2.3). Background levels of wilt were high for plants grown in compost, with 50% wilt being recorded for the control plants. Wilt symptoms developed more quickly in vermiculite (data not shown), and relatively low levels of background wilt (16% in the control) was recorded.

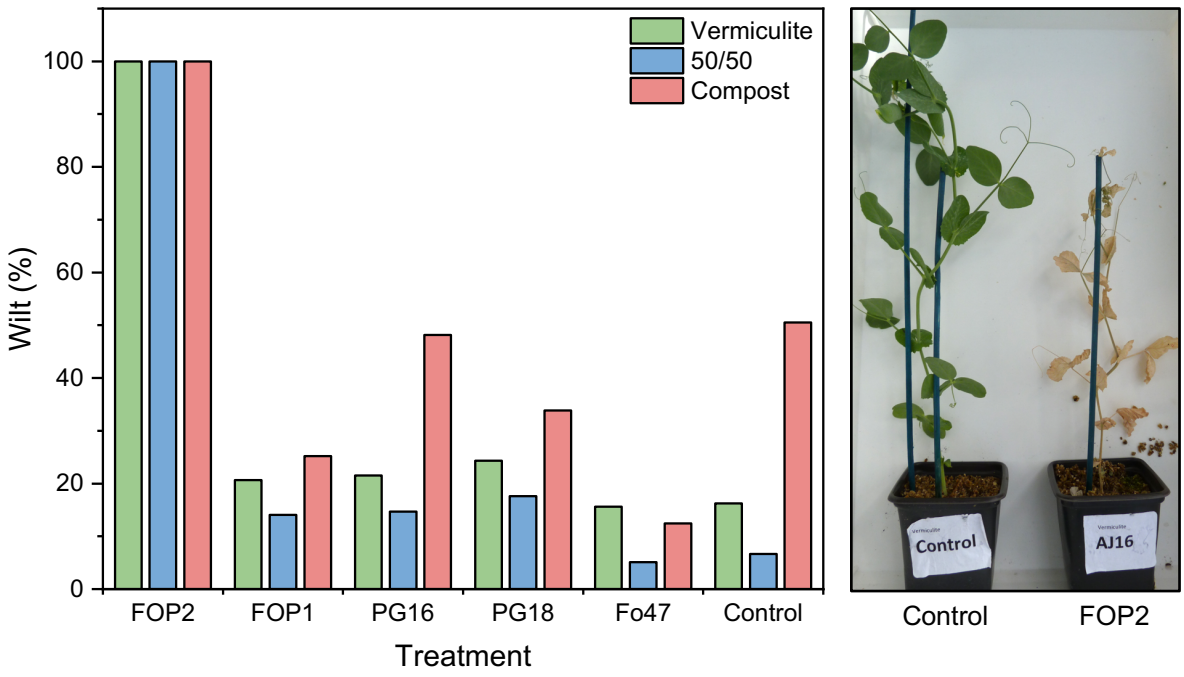


Figure 2.8 Pathogenicity of five isolates of *Fusarium oxysporum* on pea plants (cv. Avola) measured as the average percentage of wilted leaves, as part of the root dip inoculation method, 29 dpi. All data were back-transformed average proportion wilt values obtained following ANOVA analysis of logit transformed data. Image shows wilt symptoms in a pathogenic isolate (FOP2) compared to the control.

Table 2.3 Table of logit transformed ANOVA means of the proportion of wilt caused by isolates of *Fusarium oxysporum* infecting peas grown in three different growth mediums (Vermiculite, compost and a 50/50 mix of the two) at 29 dpi, using the root dip pathogenicity test. Comparisons can be made using the least significant difference (LSD) at 5% to determine significant differences between all values.

Isolate	**Logit transformed ANOVA values (% wilted leaves)**		
	Vermiculite	50/50	Compost
Control	-1.517	-2.356	0.019
Fo47	-1.566	-2.568	-1.797
FOP1	-1.254	-1.68	-1.022
FOP2	3.269	3.433	3.408
PG16	-1.214	-1.636	-0.069
PG18	-1.063	-1.442	-0.637
d.f		162	
5% LSD		0.6544	

2.4 Discussion

There have been few studies on *Fusarium* diseases in pea in the UK, despite the potential for causing yield loss. Overall this study supports the findings from surveys carried out in Europe and North America in that many *Fusarium* species were isolated from diseased peas in the field (Chittem *et al.*, 2015; Persson *et al.*, 1997). In this chapter, the *Fusarium* species isolated from diseased peas in UK fields were identified, revealing that the greatest proportion of isolates belonged to *F. oxysporum*, *F. solani* and *F. redolens*. The molecular characterisation of the *F. oxysporum* isolates derived from field samples using *TEF* gene sequencing, supported by two other housekeeping genes (*TUB2* and *RBP2*), showed that they grouped into separate phylogenetic clades compared to isolates of known FOP races, and other *formae speciales*. This could suggest that these isolates are more likely to be causing root rot in field pea rather than true Fusarium wilt. Due to visual similarities, symptoms of root rot and wilt in the field can be difficult to distinguish, especially later on in the disease process as plants with root rot start to yellow and necrose above ground (Kraft & Pfleger, 2001; Biddle & Cattlin, 2007).

All isolates collected from diseased peas in UK fields were categorised based on their morphology on PDA, before molecular identification. There was large variation in colour and aerial mycelium making identification difficult without molecular tools such as housekeeping gene sequencing. Spore size and shape was not considered, which may have reduced the number of categories with similar characteristics. Although morphological identification has traditionally been used for many years and can be reliable when used by researchers with a wealth of experience in *Fusarium* taxonomy, the development of molecular methods has allowed more reliable identification especially by those with less prior knowledge of species classification techniques. It has been noted that it can be particularly difficult to distinguish *Fusarium* species based on morphology alone as it is necessary to follow a complex identification process, logically and systematically, through all steps involved, and if certain criteria are not applied it can lead to incorrect identification (Summerell *et al.*, 2003). Other difficulties of morphological identification include the sensitivity of *Fusarium* to slight changes in environmental conditions and the profound effect this has on observed characteristics; for example, light and pH can alter spore length/production and culture pigmentation (Carlile, 1956; Avalos & Estrada, 2010; Leslie & Summerell, 2006).

Five main species of *Fusarium* were obtained from diseased pea roots in the UK in 2015, with *F. oxysporum* being the most abundant and constituting 55% of the samples collected, from fields in Yorkshire and Suffolk. This was followed by *F. solani* as the second most frequent species in 2015, with smaller numbers also observed for *F. redolens*, *F. culmorum* and *F. avenaceum*. Many of these species are common soil dwellers and may not have been the only cause of the symptoms observed. In a study of pea fields in North Dakota, a variety of *Fusarium* species were isolated, and even though *F. oxysporum* was one of the most frequent, they were found to be less virulent than the positive control isolate (*F. solani* f. sp. *pisi*). The *F. redolens* isolates they sampled were all found to be weakly pathogenic, therefore suggesting that some isolates are either contributing to disease severity through the root rot complex, or are opportunistic soil inhabitants (Chittem *et al.*, 2015). *F. graminearum*, *F. culmorum*, *F. avenaceum* and *F. poae* (all found in this study) are associated with root rot and crown rot (*Fusarium* head blight) diseases in cereals (Parry *et al.*, 1995) which are commonly included in crop rotations with pea. *F. avenaceum* has a broad host range so in crop rotations that include wheat and peas there is greater potential for inoculum levels of this pathogen to increase over time, increasing the risk of disease for peas (Feng *et al.*, 2010). In this study, *F. redolens* was also found in a relatively high proportion of samples, which is similar to findings of other pea root rot studies (Taheri *et al.*, 2011). It is also reported as a pathogen of chickpea, lentils and durum wheat, which, if grown in rotation with pea would allow spore levels to build in the soil (Taheri *et al.*, 2011). It has also been shown to be a pathogen of onion in Turkey (Bayraktar, 2010), suggesting it has a wide host range, increasing its potential to be isolated in surveys and to contribute to the root rot complex in pea. The high proportion of *F. redolens* in the samples from PGRO were mainly isolated from a root rot field trial where plants would have been exposed to additional stress (field with high root rot index), which may have allowed opportunistic infection from a generalist pathogen.

In 2016, fewer samples were collected and identified leading to fewer species (four main species) been identified. In contrast to 2015, *F. solani* was the most abundant species in 2016, followed by *F. redolens*, *F. oxysporum* and *F. avenaceum* from fields in Yorkshire and Lincolnshire. A similar study carried out recently in Canada found that the main fungi associated with root rot symptoms of field pea were *F. avenaceum*, *F. solani*, *F. acuminatum*, *F. redolens*, *F. oxysporum* and *F. tricinctum* (Taheri *et al.*, 2017). They also

showed that the principal causal agent of root rot in 2014 was *F. avenaceum* while in 2015 it was *F. solani*. Taheri *et al.* (2017) suggested that the higher than normal precipitation in 2014 would have allowed more generalist species like *F. avenaceum* (that favours waterlogged conditions) to occupy the niche usually occupied by *F. solani*, which germinates and survives in drier conditions. Feng *et al.* (2010) also found that *F. avenaceum* was the most abundant species isolated from pea roots in a wetter than normal year in central Alberta. Temperature and preceding crop may also have an impact on the abundance of species in the soil and therefore the likelihood of a more generalist pathogen being the causal agent of infection in pea (Taheri *et al.*, 2017). Differences in temperature, weather, field location, sampling time in the growing season and isolation site on the plant could all have contributed to the differences in species abundance between the two sampling years in the study reported here. It is not possible to determine from the samples collected in this study whether there was a spatial distribution of *Fusarium* species within and between fields. This was due to the small number of plants collected in each field and due to isolations being taken from stems and roots in different sampling years. In addition, the stringent selection criteria of isolates for molecular identification from each field may have artificially reduced the number of species identified.

FOP isolates from overseas were also molecularly identified using *TEF*, as even though they had supposedly been race-typed using the pea differential cultivars, some were not actually *F. oxysporum*. These had been identified using morphology alone, based on pigmentation and amount of aerial mycelium produced by each isolate (Merzoug *et al.*, 2014). As discussed earlier, this method of identification can be prone to misinterpretation. Instead, using molecular techniques such as sequencing of conserved genes similar to *TEF*, the most commonly used gene for *Fusarium* species identification, improves the accuracy and certainty of correct identification. This is because sequences can be compared to numerous characterised sequences using BLAST search tools on databases such as NCBI or FUSARIUM-ID (Geiser *et al.*, 2004). The FUSARIUM-ID database contains vouchered and well characterised sequences which correspond to publicly available cultures, which can be obtained and used for confirmation (Geiser *et al.*, 2004).

TEF gene sequencing was used in this study to identify *Fusarium* species from diseased peas in the UK and FOP isolates from overseas to understand their phylogeny. *TEF* was

able to distinguish the different *Fusarium* species into distinct clades, confirming the suitability of this housekeeping gene for molecular distinction. When characterising *F. oxysporum* isolates alone, the majority of *F. oxysporum* isolates from diseased pea plants in the UK were grouped into a single clade, containing no other sequences of FOP or other f. spp., suggesting that they originate from a different lineage to FOP. They were therefore classified as isolates causing root rot in pea and not wilt, and therefore may have evolved pathogenicity differently to FOP. The FOP isolates also grouped into distinct clades separate from isolates of other f. spp., with clades representing races 1 and 2 originating from the same branch. This could suggest that race 2 evolved from race 1 by modifying its genome in response to the deployment of race 1 resistant pea cultivars, in order to evade detection by the host. Some of the race 2 isolates also grouped into the same clade (clade 4) as the historic race 5 isolate from Warwick HRI (FOP5), suggesting these isolates could have potentially been incorrectly race typed and actually belong to race 5. The race 5 clade was separated from race 1 and 2 clades, suggesting a polyphyletic origin of FOP races, as race 5 was in a separate lineage. The polyphyletic origin of races from FOL and *F. oxysporum* f. sp. *cubense* provides evidence of horizontal gene transfer in order for races to evolve by convergent evolution (O'Donnell *et al.*, 1998; Ma *et al.*, 2010). This was discussed further in Chapter 4.

Similar clades for *F. oxysporum* and FOP were inferred from the concatenated tree of *TEF*, *TUB2* and *RPB2* sequences with the isolates from diseased peas from UK fields grouping into a distinct clade, race 1 and race 2 isolates separating into closely related but separate clades, and some of the race 2 isolates grouping in a clade with FOP5. The use of a concatenated tree of three housekeeping genes allowed a more accurate phylogeny to be inferred compared to using each gene to construct phylogenies separately (Gadagkar *et al.*, 2005). This was observed in this study as clade 3 contained FOP (e.g. FOP5, R2, F40) and *F. oxysporum* root rot isolates (e.g. PG2) in the same clade when only *TEF* was used, but with the concatenated gene sequences this clade was further characterised with most of the FOP and *F. oxysporum* isolates being separated into sub-clades.

FOP is currently controlled using resistant varieties of pea, which would keep levels of spores low in the soil, and may also help explain why FOP was not found in field samples. Varieties resistant to FOP currently grown in the UK prevent disease development, even though the pathogen is still likely to be present in the soil, but as resistance has not yet

broken down it is not of imminent concern. Pea varieties resistant to FOP are not however resistant to Fusarium root rot, and as there are currently no resistant varieties to these root rot pathogens, they will remain to be the largest problem in the UK. Difficulties in isolating FOP could also have arisen by isolating from pea roots, as it has been previously reported that isolations should always be taken from the above ground part of the plant, preferably above the fourth node (Kraft & Pfleger, 2001). In addition to root isolations, the upper stem should have also been sampled and then isolates from these two sections compared. However, Kraft and Pfleger (2001) also state that *F. solani* is rarely isolated from above ground parts of the plants, which was not the case in this study as *F. solani* was identified from stem sections between nodes 2-3. This could suggest that *F. solani* travels further up the stem into above ground parts of the plant than originally thought.

The distinction between FOP and *F. oxysporum* root rot determined using *TEF* gene sequencing was further supported by the development of pathogenicity tests to distinguish wilt and root rot symptoms. In previous research, a seed inoculation test was used to test root rot in peas (Grunwald *et al.*, 2003; Bodah *et al.*, 2016), as well as susceptibility of onions to *F. oxysporum* f. sp. *cepae* (Taylor *et al.*, 2013; Dissanayake *et al.*, 2009). In this initial test, pathogenic root rot isolates were successfully distinguished from non-inoculated (control) treatments and from the non-pathogenic isolate Fo47, where percentage emergence and in some cases subsequent seedling survival were significantly reduced. Isolates resulting in reduced seed germination suggest that they cause high levels of pre-emergence damping off, of which PG21 caused the lowest germination at <10%. In contrast, isolates resulting in the highest plant mortality indicate that they cause the most damage post emergence, of which PG18 resulted in the lowest plant survival 23 dpi. All isolates tested were from diseased peas from UK fields and all resulted in root rot symptoms at varying severities.

In the test tube pathogenicity assay all *F. oxysporum* isolates appeared highly pathogenic compared to the control. The non-pathogenic isolate Fo47 also showed significant levels of browning compared to the control, although the disease score was still significantly lower than all other treatments at 14 dpi. In contrast to the seed inoculation test, there were no significant differences between isolates. This may be because this is an artificial test using an unnatural growth medium for plants, potentially subjecting plants to additional stress. However, this approach could still be useful as an additional test

alongside the seed inoculation test as it is a rapid way to observe the symptoms of root rot pathogens.

The root dip test is widely used in studies for testing pathogenicity of FOP isolates causing vascular wilt and for determining race type (Haglund, 1989; Haglund & Kraft, 1979; Bani *et al.*, 2012; Bani *et al.*, 2014). In this initial test, different growth mediums were used to determine the optimum conditions for both the plant and pathogen, as both vermiculite and compost have been used previously (Bani *et al.*, 2012; Sharma, 2011). In all cases the FOP race 2 isolate (FOP2) resulted in 100% of leaves showing wilt symptoms at 29 dpi. PG18 resulted in low levels of wilt compared to the control, however, in the seed inoculation test it resulted in the highest plant mortality. This shows that the root dip test is suitable for displaying wilt symptoms caused by FOP isolates and not root rot symptoms caused by other *F. oxysporum* root rot causing isolates. Although the mix of 50/50 vermiculite and compost resulted in the lowest levels of background wilt (in the control), vermiculite was deemed the most suitable medium as it allowed symptoms to develop more rapidly (data not shown) and roots were much easier to remove during plant collection to assess root symptoms (data not shown), and was therefore used in subsequent tests (Chapter 5). Plants grown in compost all had high levels of background wilt (control) and it was therefore difficult to distinguish pathogenic and non-pathogenic isolates. FOP1 (FOP race 1) did not produce any significant wilt symptoms which could be due to isolates losing pathogenicity in culture (historic isolate stored for unknown amount of time), and was be investigated further in Chapter 5.

Overall, there were a range of *Fusarium* species potentially affecting peas in the UK, but their ability to cause disease still needs to be defined. Results of preliminary pathogenicity tests have helped establish different inoculation methods for potential root rot and wilt causing *F. oxysporum* isolates, which were used later in this study (Chapter 5).

3. Prediction of putative effector genes in Fusarium wilt races of FOP using whole genome sequencing

3.1 Introduction

To successfully facilitate infection, plant pathogenic fungi must be able to counteract the innate host plants immune response and manipulate the host cells physiology, which is achieved through the secretion of effector proteins (Selin *et al.*, 2016). However, they are also recognised by the host surveillance system which triggers plant defence responses, in the form of gene for gene interactions (Selin *et al.*, 2016). Infection of different host plants likely requires a different set of effectors which renders these proteins useful for distinguishing pathogenic strains, such as *formae speciales* (f. spp.) of *F. oxysporum* (van Dam *et al.*, 2016). The availability of transcriptomic and fungal genomic sequences has accelerated the numbers of effectors identified recently (Selin *et al.*, 2016), including the characterised *Secreted In Xylem* (*SIX*) genes in *F. oxysporum* f. sp. *lycopersici* (FOL), the tomato Fusarium wilt pathogen (Lievens *et al.*, 2009; Ma *et al.*, 2010; Schmidt *et al.*, 2013) (Section 4.1).

Effector prediction approaches in fungi mainly rely on relatively broad criteria, principally the presence of a secretion signal, being small in size and rich in cysteine residues (Sperschneider *et al.*, 2015). Effectors are generally secreted via the endoplasmic reticulum (ER) or Golgi route which requires a hydrophobic N-terminal signal peptide, the presence of which has been used to identify putative effectors in the secretome (Gibriel *et al.*, 2016). In addition, tools such as EffectorP improve effector prediction by identifying sequence derived properties such as length, molecular weight, protein net charge and cysteine/tryptophan content, and when combined with *in planta* expression data can predict high-priority effector candidates (Sperschneider *et al.*, 2016). Comparative studies of genomes from closely related pathogens have also helped to identify effector candidates and determine those located on core and lineage specific regions of the genome, allowing an improved understanding of effector evolution and development (Gibriel *et al.*, 2016).

Genome sequencing of the tomato infecting FOL isolate 4287 revealed a total of 15 chromosomes, compared to only 11 in *F. verticillioides* and 4 in *F. graminearum* (Ma *et al.*, 2010). Eleven of the chromosomes were syntenic between *F. oxysporum* and *F. verticillioides* and were therefore referred to as the 'core' genome. Additional unique sequences in FOL made up a significant proportion of the entire genome (40%) and were termed lineage specific (LS) regions; these included four entire chromosomes (3, 6, 14 and 15), small parts of chromosomes 1 and 2 and additional smaller unmatched contigs. These regions accounted for the larger genome size in FOL compared to the other *Fusarium* spp. (Ma *et al.*, 2010). LS regions are rich in transposable elements (TEs), transcription factors, 'secreted effectors and virulence factors' and signal transduction proteins (predicted functional classes), but lack house-keeping genes (Ma *et al.*, 2010). Among the genes identified with predicted functions were effector proteins, necrosis inducing peptides and secreted cell wall (plant/fungi) degrading enzymes, and many of these were expressed during the early stages of infection in tomato. LS regions are absent in non-pathogenic *F. oxysporum* isolates and therefore are very likely to be specifically involved in virulence of FOL (Ma *et al.*, 2010).

LS chromosome 14, also contained the sequences for all the *SIX* genes identified in FOL (Section 4.1) (Ma *et al.*, 2010; Schmidt *et al.*, 2013), suggesting they could be responsible for pathogenicity of FOL towards tomato (Ma *et al.*, 2010). *SIX* genes code for small proteins secreted into the xylem sap of tomato during infection by FOL and many have been shown to be essential for full virulence through gene knockout studies (Rep *et al.*, 2004; Houterman *et al.*, 2009; Gawehns *et al.*, 2014). Chromosome 14 was also transferred from FOL into a non-pathogenic isolate (Fo47), which then gained the ability to infect tomato plants to varying degrees, suggesting it is directly involved in pathogenicity in FOL (Ma *et al.*, 2010). It is hypothesised that LS regions in FOL may have been acquired by horizontal chromosome transfer from other *Fusarium* species or between f. spp. (Ma *et al.*, 2010). FOL is also polyphyletic comprising of at least four clonal linages, which is also a result of horizontal gene transfer (van Dam *et al.*, 2018). LS regions in FOL share very little synteny with other f. spp., such as *F. oxysporum* f. sp. *cubense* (Guo *et al.*, 2014), f. sp. *pisi* (FOP), f. sp. *ciceris*, f. sp. *medicaginis* (Williams *et al.*, 2016) and f. sp. *cepae* (FOC) (Armitage *et al.*, 2018). Although there is lack of synteny between LS regions, *SIX* genes have been identified in many other f. spp., and their presence / absence and sequences have been used as a molecular technique to

distinguish between them, which was explored further in Chapter 4 (Lievens *et al.*, 2009; van Dam *et al.*, 2016).

As mentioned previously, LS regions are rich in transposable elements (TEs), which are mobile genetic elements thought to affect genome structure and provide the plastic environment for the emergence of new virulence traits (Schmidt *et al.*, 2013). There are two main classes of TEs distinguished by their method of transposition, either via RNA or DNA. Class I TEs copy themselves into an RNA intermediate before inserting into a new site, whereas class II TEs in most cases leave the original site and integrate at a new site (Schmidt *et al.*, 2013; van Dam & Rep, 2017). Recombination between similar TEs could lead to genomic structural rearrangements, which could represent a mechanism for genetic variation in the absence of meiotic recombination in an asexual fungus like *F. oxysporum* (Schmidt *et al.*, 2013). *SIX* genes are found to be closely associated with two TEs: a miniature impala (mimp) and mFot5, which are both classes of Miniature Inverted-repeat TEs (MITEs), which are non-autonomous (unable to move on their own) and short in length (<500 bp) (Bergemann *et al.*, 2008; van Dam & Rep, 2017). Mimps are 180-220 bp long and appear to have originated from full length impala elements, part of the class II transposon superfamily, and have been shown to be active in some strains of *F. oxysporum* (Hua-Van *et al.*, 2001; Dufresne *et al.*, 2007). A total of 103 mimps were present in FOL, with the highest density found on chromosome 14, and they were present in the upstream region of all *SIX* genes, also located on chromosome 14 (Schmidt *et al.*, 2013). However, mimp deletion experiments in FOL revealed that they are not required for regulation of *SIX* gene expression *in planta*, but their consistent presence was used to identify novel effector candidates encoding proteins secreted during plant infection (Schmidt *et al.*, 2013).

FOL effector gene expression requires the presence of core-encoded conserved transcription factors such as Sge1, which has been shown to be essential for pathogenicity and required for *SIX* gene expression *in planta* (Michielse *et al.*, 2009). Thirteen predicted transcription factor genes, grouping into nine families have been identified in FOL LS regions (Schmidt *et al.*, 2013; van der Does *et al.*, 2016), with one (*FTF1*) being shown to regulate *SIX* gene expression *in planta* (Ramos *et al.*, 2007; Nino-Sanchez *et al.*, 2016). Additional predictors of pathogenicity include the presence of genes associated with carbohydrate active enzymes (CAZymes), which are involved in the degradation of plant

cell walls, an essential part of infection (Ospina-Giraldo *et al.*, 2010). A database of all current CAZymes allows users to identify CAZymes in genomic studies, and provides information on the CAZyme family and their functions (Lombard *et al.*, 2014).

The whole genome sequence analysis of a FOP race 5 isolate has also revealed the presence of a large number of transposons belonging to the *impala* family (Williams *et al.*, 2016). LS chromosomes (also known as conditionally dispensable chromosomes, CDC) have also been identified in the legume infecting *F. solani*, where a cluster of pea pathogenicity genes were found, which are involved in the detoxification of the pea defence compound pisatin (Coleman *et al.*, 2009). Although some of these genes have also been identified in FOP, they were not in the same structural cluster of pathogenicity genes as identified in *F. solani* (Coleman *et al.*, 2011). The LS chromosomes in FOP and *F. solani* have been compared, but little similarity was observed between them and similarly, FOP also had low sequence conservation when compared with FOL LS regions (Williams *et al.*, 2016). The authors concluded that there were different origins of pathogenicity within legume infecting *Fusarium* spp., and that is it a complex phenotype, not simply governed by a small set of conserved pathogenicity genes from legume specific ancestral species, or from common sequences conserved with FOL (Williams *et al.*, 2016).

The advent of long read sequencing technologies such as Pacific Biosciences and Oxford Nanopore MinION have allowed higher contig continuity and genome completeness as the long fragments generated extend across problematic or repetitive regions (Giordano *et al.*, 2017). Long read sequencing technologies are particularly useful for *de novo* genome assembly although these technologies also have a higher rate of sequencing error of 5-20%, compared to other next generation sequencing technologies (NGS, such as Illumina sequencing) at <1% (Giordano *et al.*, 2017). This high error rate requires extra correction stages in the assembly process such as Racon, Nanopolish and correction with Illumina MiSeq reads to achieve higher accuracy rates of 99.98% (Giordano *et al.*, 2017). Whole genome sequencing is an important tool for understanding the genetic differences between pathogenic and non-pathogenic isolates and the locations of pathogenicity factors in the genome.

Overall, the use of whole genome sequencing and subsequent bioinformatics analysis has enabled a detailed understanding of factors affecting pathogenicity in several *F.*

oxysporum f. spp. So far in this study a number of *F. oxysporum* isolates have been identified from UK pea fields which appear distinct in phylogeny to FOP race typed isolates from overseas.

The main aim of this chapter was to compare the genomes of different FOP races in order to investigate the genome structure and potential pathogenicity factors in FOP.

The specific objectives were:

1. To sequence and assemble genomes of three putative FOP races.
2. To determine the synteny of FOP with FOL.
3. To analyse FOP genomes and predict putative effectors.

3.2 Materials and Methods

3.2.1 FOP sample preparation and genome sequencing

Three isolates representing FOP race 1 (FOP1 EMR), 2 (F81) and 5 (R2), as identified through phylogenetic clustering (Section 2.3.2), were selected for long read whole genome sequencing (MinION). All samples were prepared and sequenced at NIAB-EMR (East Malling, Kent) using the following protocols. DNA was extracted from freeze-dried mycelium for all isolates using the Macherey-Nagel Nucleospin Plant kit (Fisher Scientific). For MiSeq libraries DNA was sheared to 550 bp fragments in a Covaris M220 sonicator and size selected using BluePippin (range 450 - 650 bp). Libraries were constructed with a PCR-free method using NEBNext End Repair, NEBNext dA-tailing and Blunt T/A ligase New England Biolabs modules and sequenced using Illumina MiSeq v3 2 x 300 bp PE. For MinION libraries, the above DNA was size selected using BluePippin for fragments larger than 8 kb using high pass filtering. Libraries were constructed using the Ligation Sequencing Kit 1D and sequenced on a gridION with local basecalling according to the manufacturer's protocol (Oxford Nanopore).

3.2.2 FOP genome assembly and gene prediction

All bioinformatics work including genome assembly, analysis and functional annotation was carried out with the aid of Dr Andrew Armitage (NIAB-EMR, Kent). MinION reads for each of the three FOP isolates (FOP1 EMR, F81 and R2) were assembled using SMARTdenovo (Ruan, 2016), an assembly package specifically designed for long read sequencing. Assemblies were corrected using Racon for ten rounds to perform initial correction of the assembly using MinION fastq files (Vaser *et al.*, 2017), followed by Nanopolish (Loman *et al.*, 2015) which corrected reads based on the raw signal data measured on the nanopore sequencer, and finally polished using Illumina MiSeq reads in Pilon to correct remaining mis-assemblies (Walker *et al.*, 2014). Quast was used to summarise assembly statistics and BUSCO v3 was used to assess completeness of gene space within the assembly (Gurevich *et al.*, 2013; Simão *et al.*, 2015). RepeatModeler and transposonPSI were used to identify repetitive and low complexity regions (Smit *et al.*, 2013-2015; Haas, 2007).

RNAseq data for each isolate (FOP1 EMR, F81 and R2) derived from infected pea roots (one of the three biological replicates, Chapter 4) was used to provide evidence for gene prediction by aligning reads to the corresponding *de novo* assembled genomes using STAR (Dobin *et al.*, 2013). Gene prediction was performed on softmasked genomes using BRAKER1 v.2 (Hoff *et al.*, 2016), which is a program that trains AUGUSTUS 3.1 gene prediction (Stanke & Morgenstern, 2005). Separately, RNAseq reads were also assembled into transcripts using Cufflinks (Trapnell *et al.*, 2010), before being used to predict additional genes using CodingQuarry v.2, that may have been missed previously using BRAKER1 (Testa *et al.*, 2015). Gene models from BRAKER and CodingQuarry were combined to provide a number of predicted genes per isolate; however, CodingQuarry genes were discarded if there was also a BRAKER gene model predicted in the same location. Gene prediction accuracy was confirmed using BUSCO v3 (Simão *et al.*, 2015).

3.2.2.1 Functional annotation and effector prediction

InterProScan-5 was used to functionally characterise predicted protein sequences from the generated gene models by comparing identifiable features to known protein families, domains and functional sites (Jones *et al.*, 2014). In addition, Swiss-Prot (2016 release) was also used to identify proteins and provide a high level of annotation such as function and posttranslational modifications (Bairoch & Apweiler, 2000). Gene models were also compared to PHIbase, a database of pathogenicity and effector genes from fungi and other pathogens which infect plants, animals and insects, to annotate genes potentially involved in pathogen-host interactions in FOP (Urban *et al.*, 2017).

Putative effectors were predicted using different tools. EffectorP (v.1.0) was used to identify fungal effectors using features such as sequence length, molecular weight, protein net charge and cysteine, serine and tryptophan content (Sperschneider *et al.*, 2016). Putative secreted proteins were predicted from gene models via the presence of a signal peptide using SignalP v.4.1 (Petersen *et al.*, 2011). In addition, proteins predicted to contain transmembrane domains were identified using TMHMM v.2.0 (Krogh *et al.*, 2001) and were excluded from lists of signal peptide containing proteins, as these are unlikely to represent cytoplasmic or apoplastic effectors.

Predicted secreted proteins were screened for carbohydrate active enzymes (CAZymes) using HMM models from the CAZy database (Lombard *et al.*, 2014) and HMMER3, a tool used for searching sequences against a database of trained HMM models (Finn *et al.*, 2011). In addition, genes with transcription factor annotations were identified using the DBD:Transcription factor prediction database (Wilson *et al.*, 2008) and a further set of InterProScan annotations (Shelest, 2017). BLAST searches for homologs to *Fusarium* transcription factors (*FTF*) from Nino-Sanchez *et al.* (2016) were conducted against the FOP genomes, and then intersected with gene models. Similarly, sequences of previously characterised *Secreted In Xylem* (*SIX*) genes in FOL (obtained from NCBI) were used in BLAST searches against the FOP genome assemblies and also against the gene models.

As mentioned in the introduction (Section 3.1), mimp sequences are important in identifying potential pathogenicity genes (Schmidt *et al.*, 2013). Mimp sequences were identified by searching for the sequence of the mimp 3' inverted repeat (TT[TA]TTGCNNCCCACTG) (Schmidt *et al.*, 2013). Genes that were predicted as secreted and within 2 kb of a mimp sequence were marked as candidates for being under the influence of a mimp-containing promoter.

3.2.3 Synteny of FOP and FOL genomes using orthologous genes

Genes in orthogroups containing a single gene from each FOP isolate (FOP1 EMR, F81 and R2) and from FOL were used to assess the synteny between FOP and FOL genomes. Only genes that had a 1:1 relationship between a FOP isolate and FOL were used to draw links between FOP contigs and previously characterised FOL chromosomes (Ma *et al.*, 2010). These relationships were visualised with Circos plots (Krzywinski *et al.*, 2009), and FOP contigs were re-arranged to match the size and order of the FOL chromosomes, to try and infer the core and lineage specific chromosomes for each FOP isolate.

3.2.4 Position of putative effectors in FOP genomes

The location, product size, contig number and other functional annotations of the *SIX* genes found in the genomes of the three FOP isolates (FOP1 EMR, F81 and R2) were also examined and compared between putative races. The location of predicted effector genes (EffectorP, CAZymes, mimps and *SIX* genes) were annotated on contigs from each genome and visualised in size order using Circos plots (Krzywinski *et al.*, 2009).

3.2.5 Orthology analysis of FOP isolates with FOC and FOL

Identifying orthology between proteins is a valuable tool in gene annotation as they are likely to retain function over time (Li *et al.*, 2003). Ortholog group analysis was carried out between predicted proteins identified from each of the FOP genomes (FOP1 EMR, F81 and R2), *F. oxysporum* f. sp. *cepae* (FOC) isolate FUS2 (Armitage *et al.*, 2018), and FOL isolate 4287 (Ma *et al.*, 2010; Armitage *et al.*, 2018). OrthoMCL v.2.0.9 was used to compare the predicted proteins from these genomes (Li *et al.*, 2003) and Venn diagrams visualising common orthogroups were plotted using the R package VennDiagram (Chen & Boutros, 2011).

Orthology was used to examine shared genes between the three FOP isolates, and were filtered if they were predicted as secreted. Unique genes to each race were also filtered based on whether they were predicted as secreted, and both were used to predict putative effectors.

3.3 Results

3.3.1 Genome assembly and gene prediction of three FOP genomes

All analyses included in this chapter were carried out with the aid of Dr Andrew Armitage (NIAB-EMR, Kent). The assembly of the FOP genomes using MinION and MiSeq sequence data produced assembly sizes of 66 Mb (FOP1 EMR), 60 Mb (F81) and 56 Mb (R2) which were comparable to the FOC and FOL reference genomes of 53 Mb and 62 Mb respectively (Armitage *et al.*, 2018) (Table 3.1). The smallest genome, R2 (race 5), was assembled into 33 contigs, compared to 82 and 74 contigs for FOP1 EMR (race1), and F81 (race 2) respectively. Gene space within all the assemblies was shown to be comparable with over 99% of the 3725 core Sordariomycete genes (BUSCO) present, which is comparable to the reference genomes of FOC and FOL. The genome of FOP1 EMR had the highest percentage of bases repeatmasked, at 24%, compared to the other genomes at only 14% (F81), 11% (R2), 11% (FOC) and 16% (FOL).

Gene prediction using Braker and CodingQuarry resulted in 19275 – 22115 genes across the FOP isolates (Table 3.1). BUSCO analysis showed a low false negative rate with 98-99% of Sordariomycete genes predicted. Predicted proteins were tested for the presence of signal peptides, which implies the protein is secreted, and 1538 were identified in F81, with slightly lower numbers in FOP1 EMR (1385) and R2 (1487). These results were comparable to the number of secreted proteins in FOC and FOL genomes (Table 3.1).

Table 3.1 Summary statistics for the assembled genomes and predicted gene models for three *Fusarium oxysporum* f. sp. *pisi* (FOP) isolates FOP1 EMR (race 1), F81 (race 2) and R2 (race 5). Reference values are included for genomes of *F. oxysporum* f. sp. *cepae* (FOC) and *F. oxysporum* f. sp. *lycopersici* (FOL).

	FOP			FOC	FOL
	FOP1 EMR	**F81**	**R2**	**Fus2***	**4287****
Assembly stats					
Total coverage (fold)	62	44	47	214	6.8
Sequencing technology	MinION + MiSeq	MinION + MiSeq	MinION + MiSeq	PacBio + MiSeq	Sanger (WGS)
Assembly size (Mb)	66	60	56	53	62
Contigs	82	74	33	34	15 + 73
Largest contig (kb)	6777	4422	6565	6434	6855
N50 (kb)	2900	2659	3457	414	4590
Sordariomycete genes (BUSCO)	3687	3686	3681	3687	3599
% Sordariomycete genes (BUSCO)	99	99	99	99	97
% Repeatmasked	24	14	11	11	16
Gene prediction					
Total genes	22115	20484	19275	18855	20925
Total proteins	22273	20699	19477	19371	27347
Sordariomycete genes (BUSCO)	3668	3673	3666	3668	3577
% Sordariomycete genes (BUSCO)	98	99	98	98	96
Secreted Genes	1385	1538	1487	1449	1493

*FOC Fus2 – *Fusarium oxysporum* f. sp. *cepae* (isolate Fus2), Armitage *et al.* (2018)
**FOL 4287 – *Fusarium oxysporum* f. sp. *lycopersici* (isolate 4287), Ma *et al.* (2010)

Table 3.2 Candidate effector genes in *Fusarium oxysporum* f. sp. *pisi* (FOP) isolates FOP1 EMR (race 1), F81 (race 2) and R2 (race 5), including EffectorP, CAZymes and *Secreted In Xylem* (*SIX*) genes. Secreted genes identified within 2 kb of a mimp and transcription factors are also included.

	FOP1 EMR	**F81**	**R2**
Effector candidates			
Total EffectorP genes predicted	3750	3406	3188
Secreted and EffectorP	340	341	336
Total CAZymes	861	986	972
Secreted and CAZymes	343	418	412
Number of *SIX* gene homologs	8	5	3
Mimps			
Mimps in genome	203	126	142
Genes within 2 kb of a mimp	142	101	121
Secreted genes within 2 kb of a mimp	24	17	21
Transcription factors			
Genes with transcription factor annotations	1895	1664	1590
FTF hits in gene models	6	3	3

Effector prediction using EffectorP identified 3750 candidate effector genes in FOP1 EMR (race 1), with 3406 identified in F81 (race 2) and 3188 in R2 (race 5). However, when these results were combined with those predicted as secreted (Table 3.1) only 336 – 341 were predicted as putative secreted effectors (Table 3.2). Carbohydrate active enzymes were also found in all three genomes, with the number secreted being similar to those predicted by EffectorP. The F81 genome contained the most CAZymes (418), with similar numbers identified in R2 (412) and slightly less in FOP1 EMR (342). BLAST searches for the 14 *SIX* genes (from FOL) against the three genomes identified eight in FOP1 EMR (race 1), five in F81 (race 2) and three in R2 (race 5) indicating a difference in *SIX* gene complement between FOP races. As *SIX* genes in FOL were all identified in proximity to a mimp (3.1 Introduction, Schmidt *et al*., 2013), BLAST searches (nucleotide) for mimp sequences were carried out against the FOP genomes, revealing 203 in FOP1 EMR, 142 in R2 and 126 in F81. These were intersected with predicted secreted effectors from the gene models to identity putative secreted genes within 2 kb of a mimp. This approach identified 24 (FOP1 EMR), 17 (F81) and 21 (R2) putative secreted candidate effectors in the three FOP genomes that were in close proximity to a mimp (Table 3.2). Transcription factors which are essential regulators of gene expression (1590-1895) were also identified in the FOP genomes (Table 3.2), and *Fusarium* specific transcription factors (*FTF*) were also identified resulting in six hits in the gene models for FOP1 EMR, and three for both F81 and R2.

3.3.2 Synteny of FOP genome with FOL genome using orthologous genes

The FOL genome (Ma *et al.*, 2010) was used as a reference genome to map contigs from the three FOP genomes (FOP1 EMR, F81 and R2) in order to identify core chromosomes through shared genes in a 1:1 relationship.

Contigs in all three FOP genomes were grouped into 11 potential core chromosomes based on synteny with the 11 FOL core chromosomes. The remaining contigs from all the genomes were classed as ungrouped and therefore named as FOP lineage specific (LS) regions as they shared few orthologous genes with FOL (Figure 3.1, Figure 3.2, Figure 3.3). The LS whole chromosomes (14, 3, 6, 15) in the FOL genome share very few orthologous genes with FOP1 EMR (race 1), F81 (race 2) and R2 (race 5) as expected. In the FOP1 EMR genome, there was a small section of mis-matched contigs where location of orthologous genes crossed over between contigs three and two in FOP1 EMR with chromosomes two and four in FOL (Figure 3.1). This also occurred in different contigs in F81 (Figure 3.2) and R2 (Figure 3.3) when compared with FOL, but to less of an extent. The number of contigs constituting an inferred chromosome varies across the isolates, with examples of whole chromosomes being sequenced (one contig) and chromosomes being inferred from multiple scaffolded contigs.

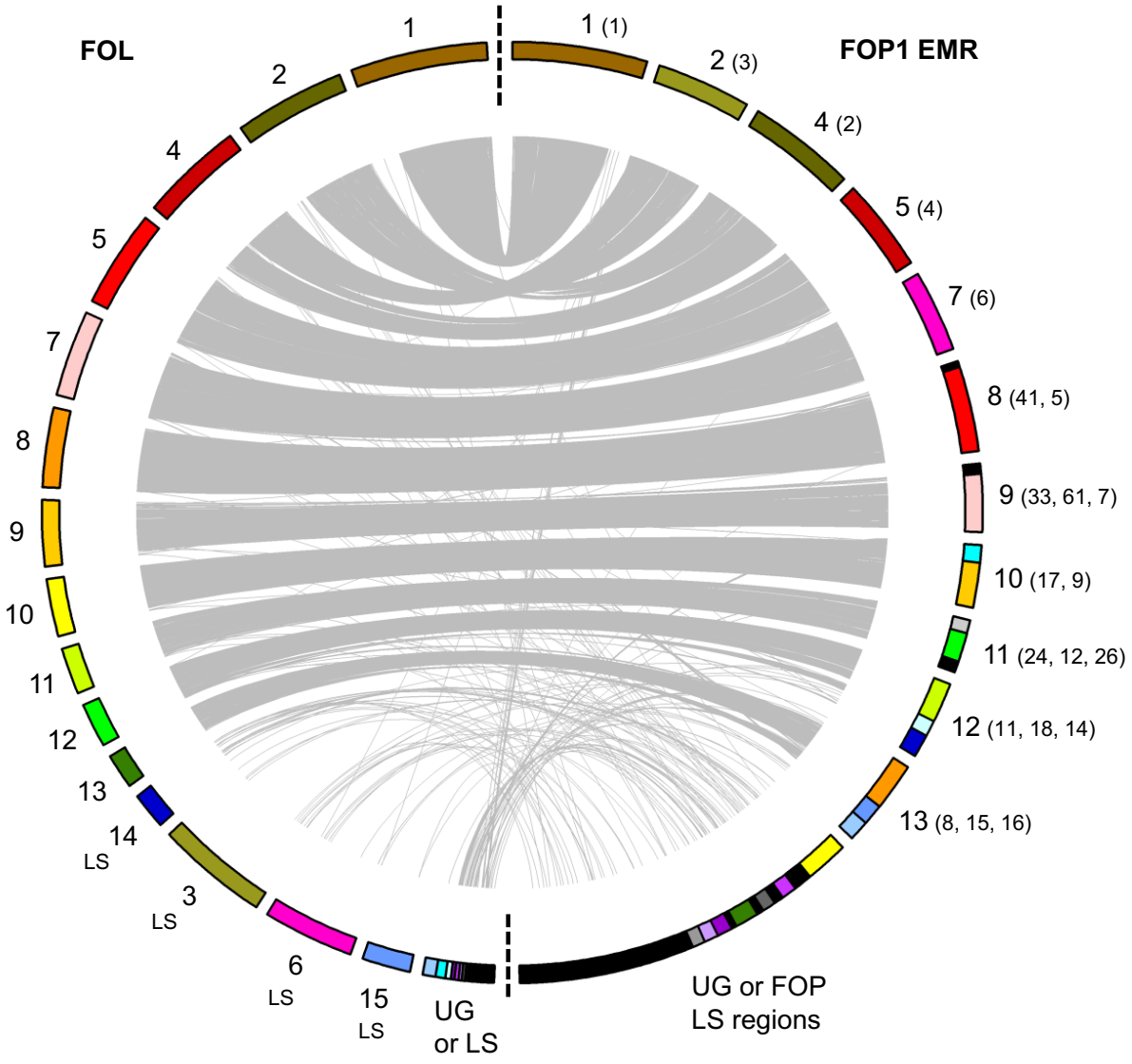


Figure 3.1 Synteny of *Fusarium oxysporum* f. sp. *pisi* (FOP) isolate FOP1 EMR (race 1) contigs with chromosomes from the published *F. oxysporum* f. sp. *lycopersici* (FOL) isolate 4287. Relationships are through linking single copy orthologous genes present in both genomes. Core chromosomes in FOL show high synteny with contigs in FOP1 EMR, which were arranged into the corresponding core chromosomes labelled with the chromosome number and the contig numbers in brackets. Ungrouped (UG) contigs in FOP1 EMR were designated as lineage specific (LS) regions.

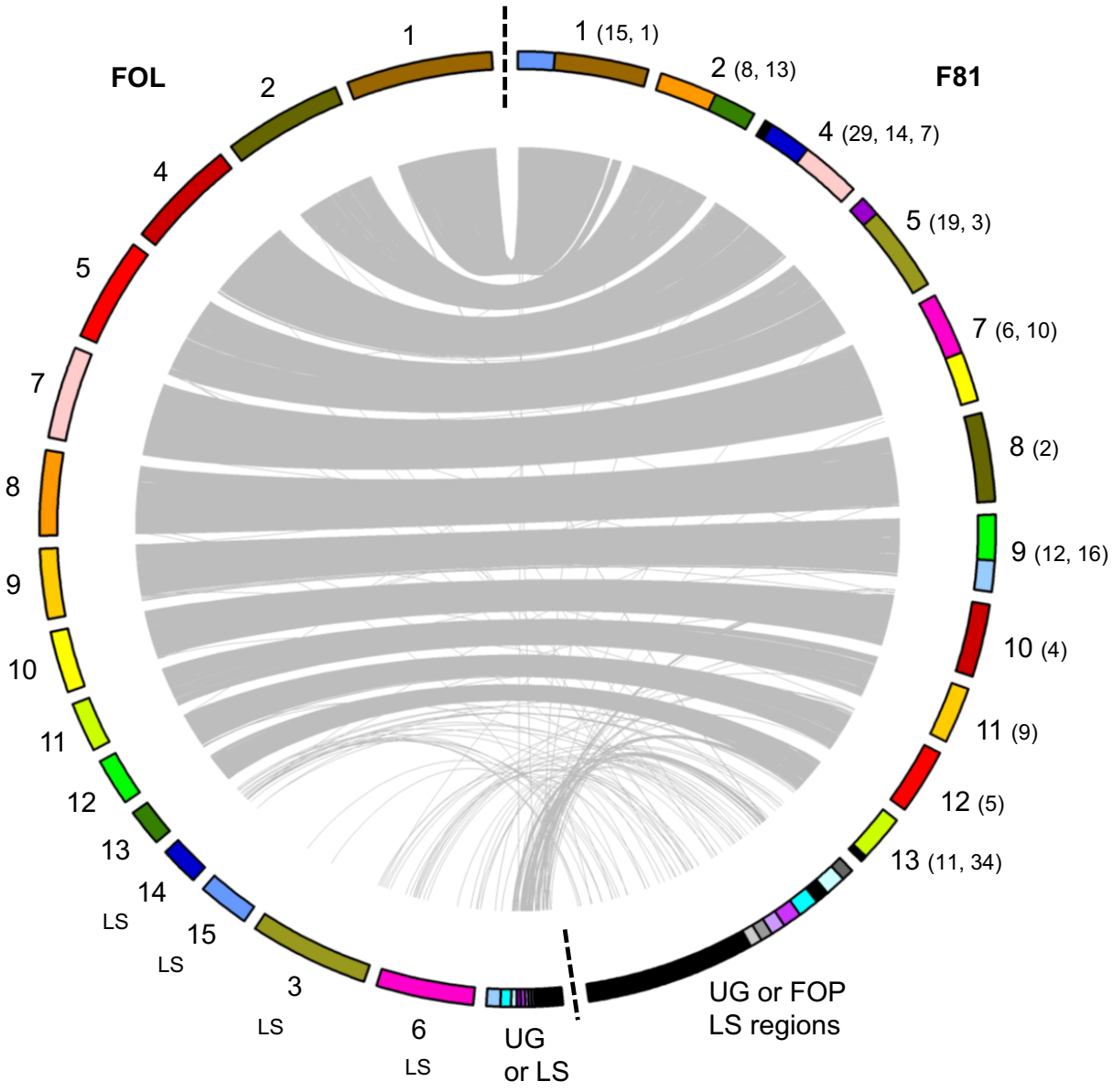


Figure 3.2 Synteny of *Fusarium oxysporum* f. sp. *pisi* (FOP) isolate F81 (race 2) contigs with chromosomes from the published *F. oxysporum* f. sp. *lycopersici* (FOL) isolate 4287. Relationships are through linking single copy orthologous genes present in both genomes. Core chromosomes in FOL show high synteny with contigs in F81, which were arranged into the corresponding core chromosomes labelled with the chromosome number and the contig numbers in brackets. Ungrouped (UG) contigs in F81 were designated as lineage specific (LS) regions.

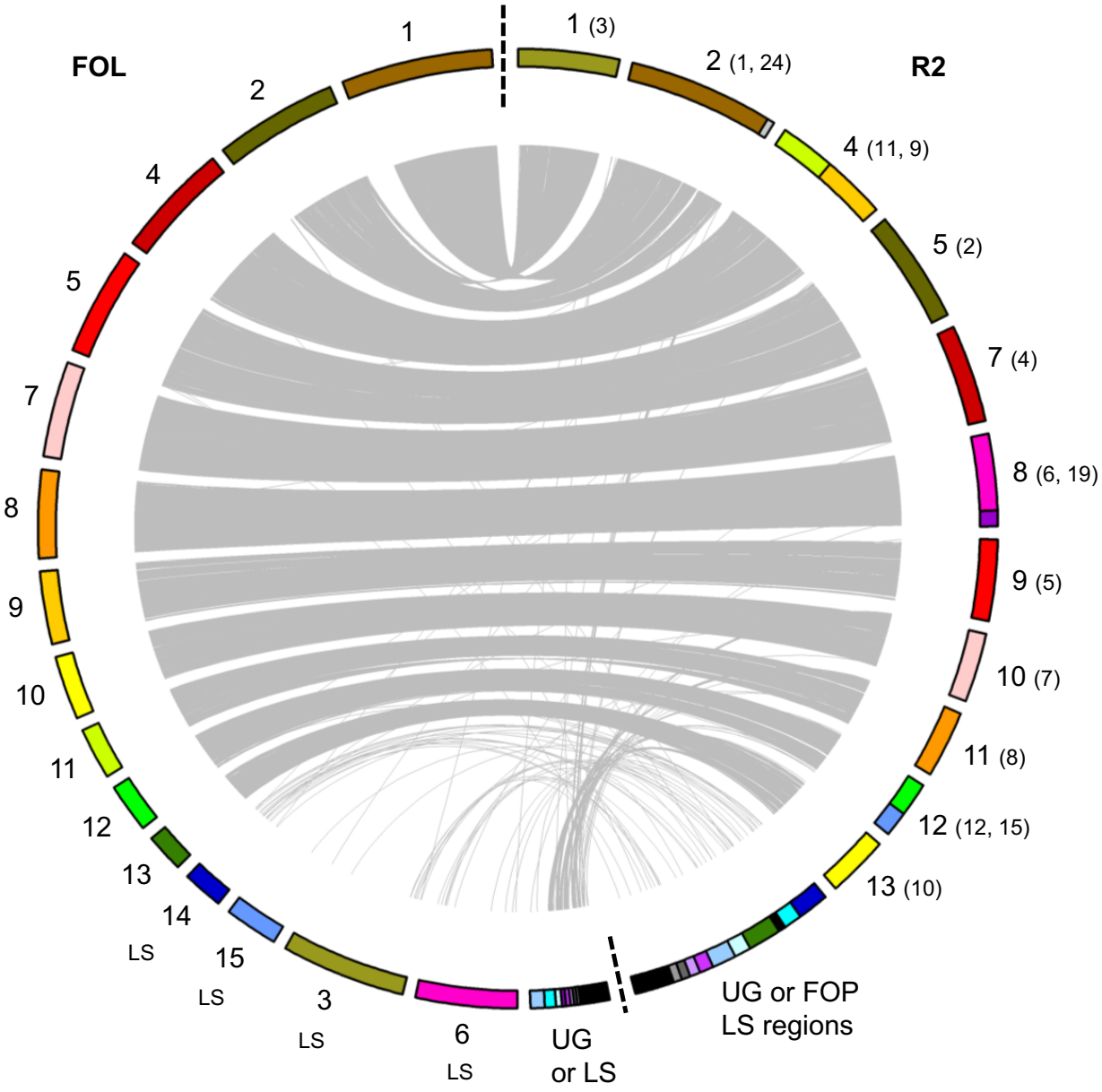


Figure 3.3 Synteny of *Fusarium oxysporum* f. sp. *pisi* (FOP) isolate R2 (race 5) contigs with chromosomes from the published *F. oxysporum* f. sp. *lycopersici* (FOL) isolate 4287. Relationships are through linking single copy orthologous genes present in both genomes. Core chromosomes in FOL show high synteny with contigs in R2, which were arranged into the corresponding core chromosomes labelled with the chromosome number and the contig numbers in brackets. Ungrouped (UG) contigs in R2 were designated as lineage specific (LS) regions.

3.3.3 Positions of putative effectors in FOP genomes

The locations of putative FOP effectors were visualised on circos plots. Contigs were arranged in size order and secreted EffectorP results, secreted CAZymes, mimps and *SIX* genes were mapped. Based on the synteny of genomes with FOL (above), contigs were grouped into core chromosomes and LS (or ungrouped) regions. In FOP1 EMR, 79% of putative effectors predicted by Effector P (blue) were located on the core chromosomes, with more than 20 genes on each of contigs 1, 2, 5, 6, 8 and 11 (Figure 3.4). Contig 11 had the greatest density of EffectorP putative effectors (based on contig size). In addition, 91% of the predicted secreted CAZymes were also located on the core chromosomes with the highest number (42 out of the 343 secreted CAZymes) located on contig 11 (Figure 3.4). Therefore, the FOP LS regions contained few predicted secreted EffectorP and CAZymes. However, LS regions in FOP1 EMR contained 82% of the mimps identified in the genome, and all of the *SIX* genes (Figure 3.4.)

In the F81 genome, 78% of putative effectors predicted by EffectorP and 87% of predicted secreted CAZymes were found on the contigs constituting the core chromosomes, with the highest numbers located on contigs 5, 9 and 11 (Figure 3.5). As with FOP1 EMR there were very few predicted secreted EffectorP and CAZymes on the LS regions in the F81 genome. However, 89% of the identified mimps were located on the LS regions of the genome, in addition to all but one (*SIX9*) of the *SIX* genes.

The R2 (race 5) assembly comprised 33 contigs, less than those in FOP1 EMR (race 1) and F81 (race 2) (Table 3.1; Figure 3.6). There was a total of 336 putative secreted effectors (Table 3.1) predicted by EffectorP, with 88% of these located on core chromosomes (Figure 3.6). Additionally, 91% of predicted secreted CAZymes were also located on these core chromosomes. The highest density (visually) of EffectorP and CAZymes were located on contigs 8, 12 and 15, considering the sizes of the contigs. The LS regions contained 92% of identified mimps, in addition to all of the identified *SIX* genes.

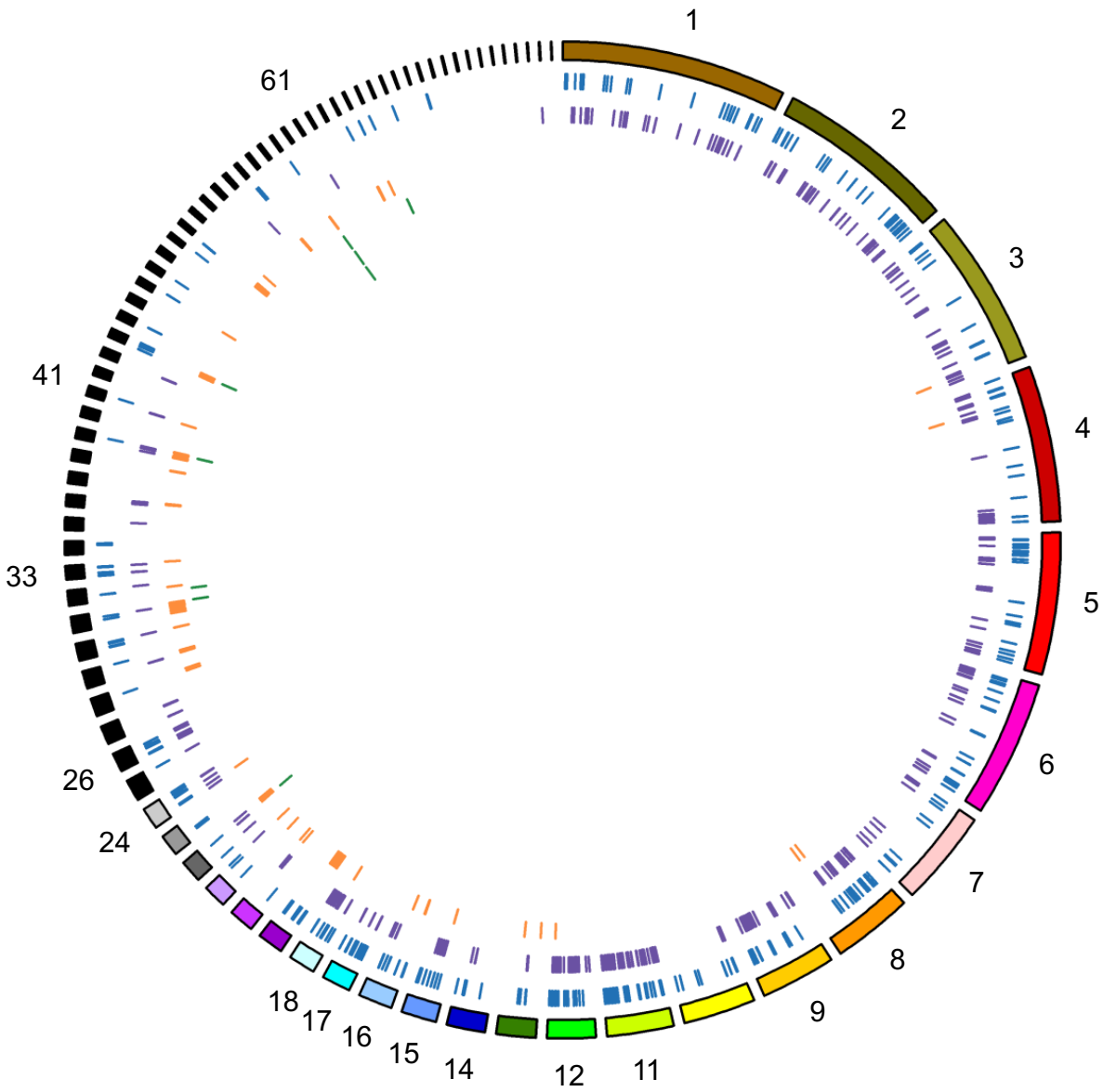


Figure 3.4 Individual contigs of *Fusarium oxysporum* f. sp. *pisi* isolate FOP1 EMR (race 1) in size order. Contigs are labelled with their contig number if they form part of the core genome. Contigs are annotated with genes identified as putative effectors by EffectorP (blue), CAZymes (purple), mimps (orange) and *Secreted In Xylem* (*SIX*) genes (green).

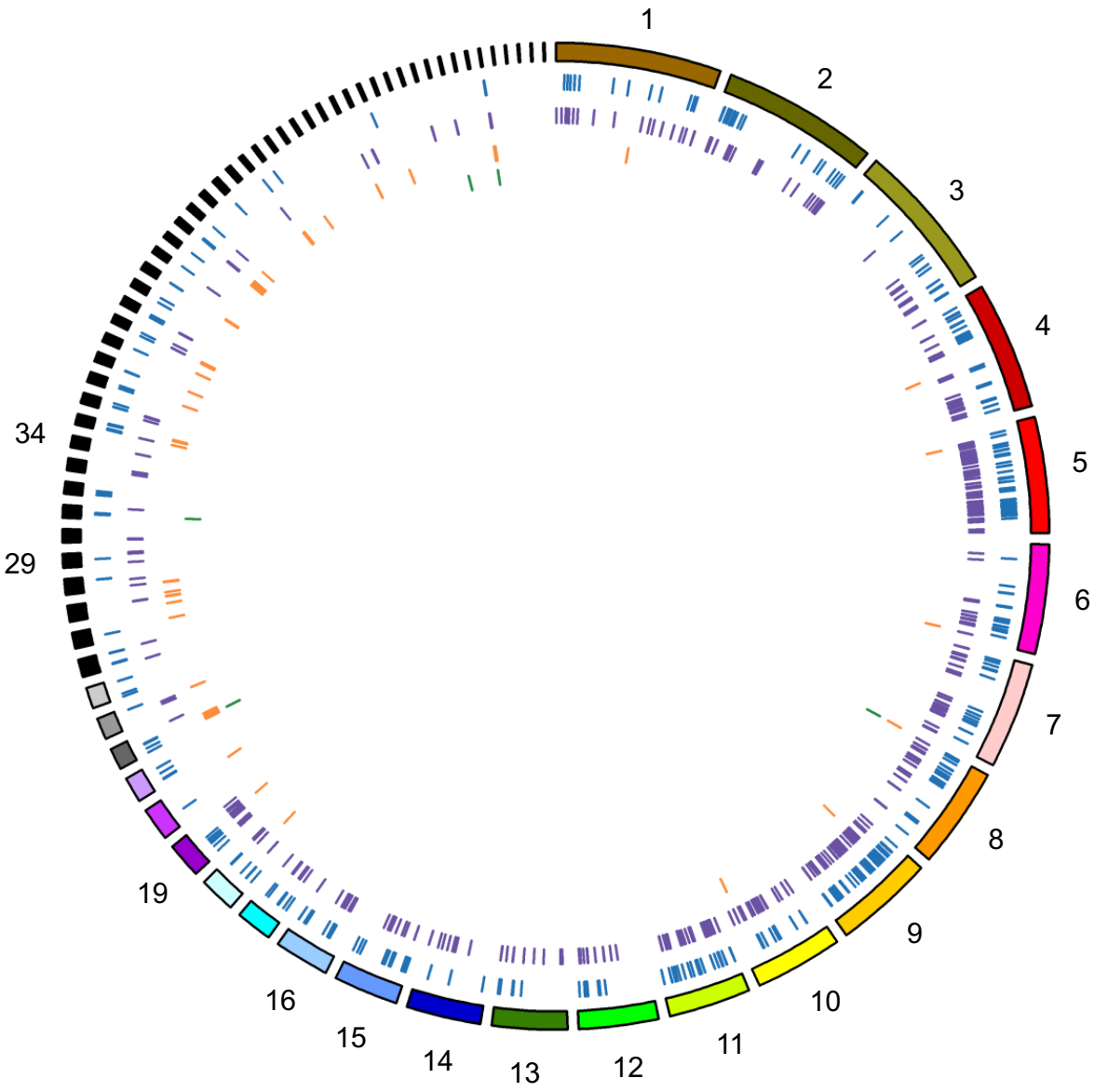


Figure 3.5 Individual contigs of *Fusarium oxysporum* f. sp. *pisi* isolate F81 (race 2) in size order. Contigs are labelled with their contig number if they form part of the core genome. Contigs are annotated with genes identified as putative effectors by EffectorP (blue), CAZymes (purple), mimps (orange) and *Secreted In Xylem* (*SIX*) genes (green).

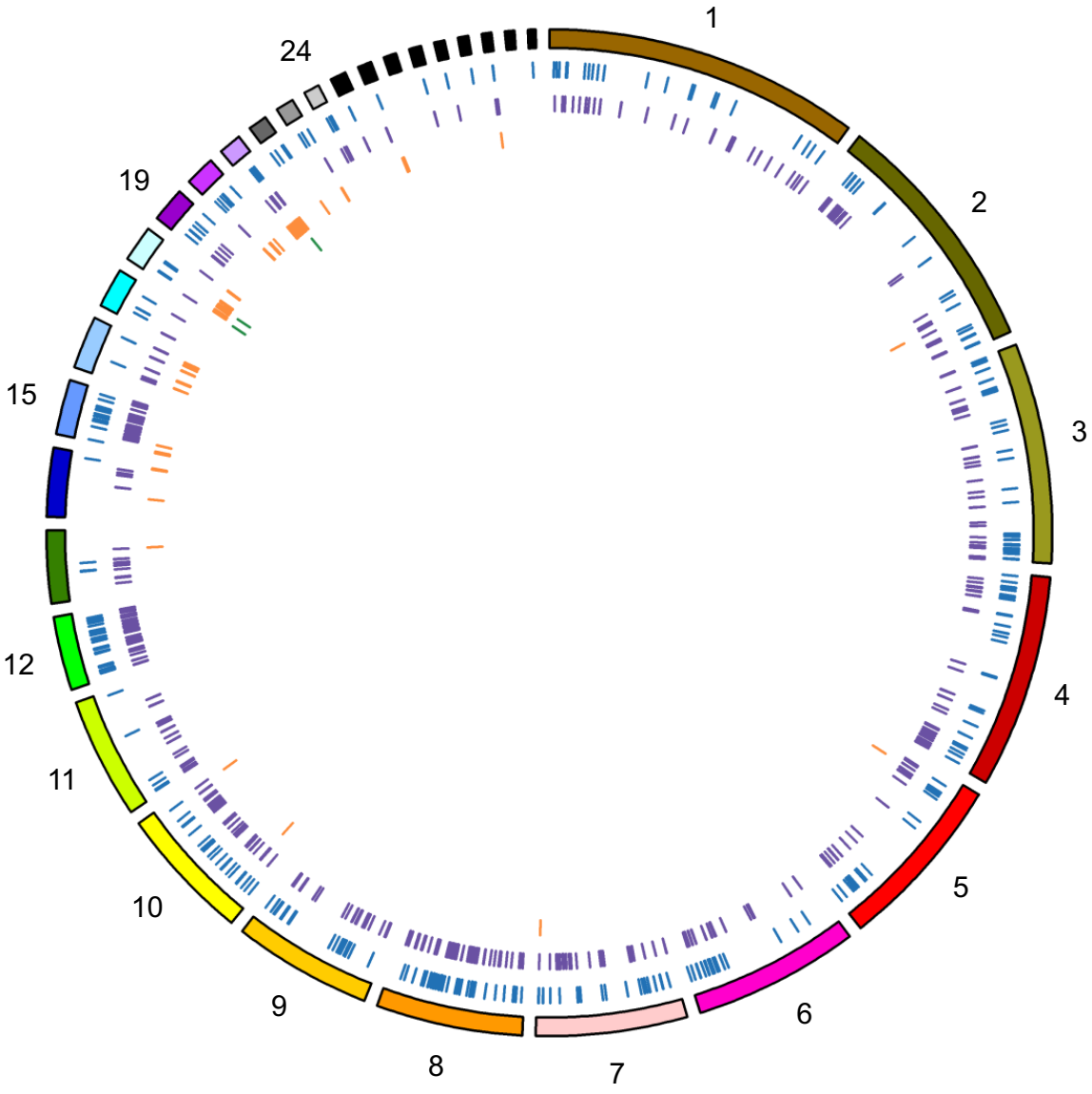


Figure 3.6 Individual contigs of *Fusarium oxysporum* f. sp. *pisi* isolate R2 (race 5) in size order. Contigs are labelled with their contig number if they form part of the core genome. Contigs are annotated with genes identified as putative effectors by EffectorP (blue), CAZymes (purple), mimps (orange) and *Secreted In Xylem* (*SIX*) genes (green).

SIX genes have been shown to be important in pathogenicity of FOL against tomato plants, and were all found to be located on LS chromosomes (3.1 Introduction). Therefore, the presence of *SIX* gene homologs and their locations in the FOP genomes (Table 3.3) was determined. For FOP1 EMR, eight *SIX* genes *SIX1*, *SIX6*, *SIX7*, *SIX9*, *SIX10*, *SIX11*, *SIX12* and *SIX14* (with a second copy of *SIX6*) were identified in BLAST searches of the whole genome, however *SIX9* and *SIX14* were not identified as predicted genes in the gene models from Braker and CodingQuarry (Table 3.1). As mentioned above, the *SIX* genes were located on the LS regions, which is comparable to the locations of *SIX* genes in FOL (3.3.2). Three of the *SIX* genes were located within 2 kb of a mimp (*SIX7*, *SIX10* and *SIX12*), a common feature of *SIX* genes in FOL (3.1 Introduction, Schmidt *et al*., 2013), and all but *SIX7* (not known for *SIX9* and *SIX14* as not identified in gene models) were predicted as putative effectors by EffectorP (Table 3.3).

In F81, five *SIX* genes (*SIX1*, *SIX6*, *SIX9*, *SIX13* and *SIX14*) were identified in the genome (with an additional copy of *SIX1*) with all but *SIX13* predicted as being secreted (Table 3.3). *SIX1*, *SIX6* and *SIX13* were located within 2 kb of a mimp, and *SIX6*, *SIX9* and *SIX14* were predicted as putative effectors by EffectorP. All *SIX* genes, other than *SIX9*, were located on LS regions in the genome when compared to FOL.

In the R2 genome, three *SIX* genes were identified, *SIX1*, *SIX6* and *SIX13* (with 2 additional copies of *SIX1*), with all of them being predicted as secreted. All three copies of *SIX1* were located on the same contig (18), and two out of three were located within 2 kb of a mimp, as with all other *SIX* genes (Table 3.3). All *SIX* genes were located on LS regions in R2, as with the *SIX* genes in other races. *SIX6* was the only *SIX* gene predicted as a putative effector by EffectorP for R2.

The different complement of *SIX* genes identified in the three FOP isolates of different putative races (1, 2 and 5) have enabled a race specific pattern to be identified. It is therefore hypothesised that race 1 isolates (based on FOP1 EMR) contain eight *SIX* genes, race 2 isolates (based on F81) contain five *SIX* genes, and race 5 isolates (based on R2) contain three *SIX* genes. This was investigated further in Chapter 4.

Table 3.3 Presence and location of *Secreted In Xylem* (*SIX*) gene homologs in the whole genomes of three *Fusarium oxysporum* f. sp. *pisi* (FOP) isolates (FOP1 EMR, F81 and R2). Additional annotation of gene size, secretion status, location in relation to a mimp and EffectorP status is included. * indicates genes which were identified in the whole genomes but not in the predicted gene models using Braker and CodingQuarry.

Isolate	*SIX* gene	Gene	Contig	Location	Size (bp)	Secreted	2 kb of a mimp	EffP
FOP1 EMR	*SIX1*	g20529	39	LS	344			Yes
FOP1 EMR	*SIX6*	g17863	22	LS	726	Yes		Yes
FOP1 EMR	*SIX6*	g20948	44	LS	464	Yes		Yes
FOP1 EMR	*SIX7*	g21760	59	LS	636		Yes	
FOP1 EMR	*SIX10*	g20977	44	LS	519	Yes	Yes	Yes
FOP1 EMR	*SIX11*	g21938	66	LS	335	Yes		Yes
FOP1 EMR	*SIX12*	g21759	59	LS	431	Yes	Yes	Yes
FOP1 EMR	*SIX9**		32	LS	290	UN	UN	UN
FOP1 EMR	*SIX14**		31	LS	140	UN	UN	UN
F81	*SIX1*	g20369	66	LS	848	Yes		
F81	*SIX1*	g18035	28	LS	848	Yes	Yes	
F81	*SIX6*	g20417	69	LS	842	Yes	Yes	Yes
F81	*SIX9*	g8636	8	Core	726	Yes		Yes
F81	*SIX13*	g17065	22	LS	356		Yes	
F81	*SIX14*	g18445	31	LS	431	Yes		Yes
R2	*SIX1*	g17379	18	LS	316	Yes		
R2	*SIX1*	g17417	18	LS	842	Yes	Yes	
R2	*SIX1*	g17623	18	LS	848	Yes	Yes	
R2	*SIX6*	g18204	21	LS	842	Yes	Yes	Yes
R2	*SIX13*	g17466	18	LS	842	Yes	Yes	

UN = Unknown

3.3.4 Orthology analysis of FOP isolates with FOC and FOL

Orthology analysis clustered proteins from FOP isolates FOP1 EMR (race 1), F81 (race 2) and R2 (race 5), as well as FOC and FOL into 18986 orthogroups. Of these, 10592 orthogroups were shared between all of the isolates (Figure 3.7), which represented 89692 shared proteins. FOL contained the largest number of unique orthogroups with 2353, compared to FOP1 EMR (769), F81 (724) and R2 (441) which all had fewer unique orthogroups (Figure 3.7). There were 133 orthogroups shared between FOP1 EMR, F81 and R2, which is only marginally higher than between the three FOP isolates and FOL, at 110 orthogroups. There were more than double the number of shared orthogroups between F81 and R2 at 116, compared to FOP1 EMR and R2 at 51 orthogroups. This was considerably fewer than those shared between FOP1 EMR and F81 with 278 orthogroups.

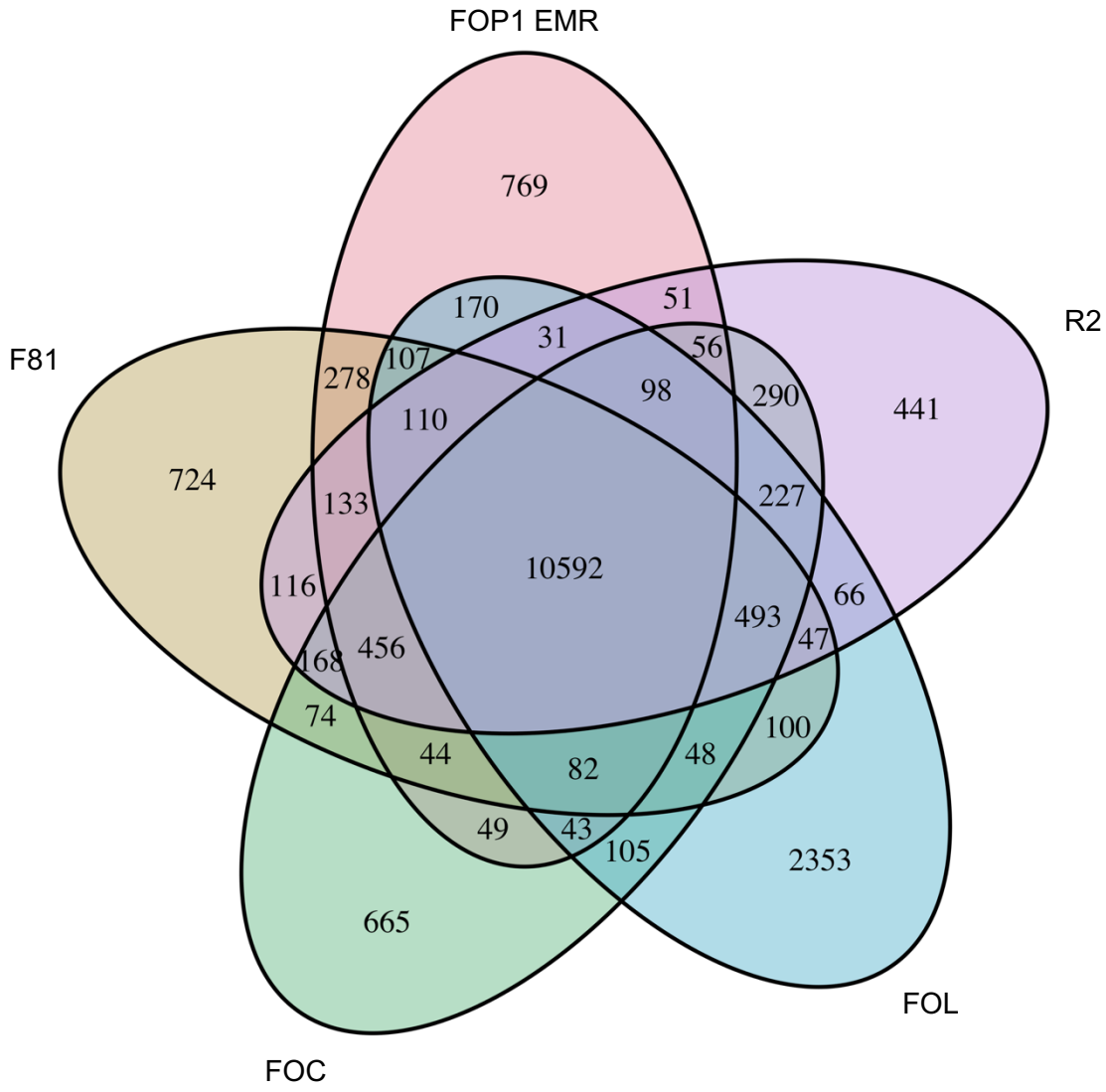


Figure 3.7 Shared and unique orthogroups between three *Fusarium oxysporum* f. sp. *pisi* (FOP) isolates FOP1 EMR (race 1), F81 (race 2) and R2 (race 5), as well as two reference genomes *F. oxysporum* f. sp. *cepae* (FOC) and *F. oxysporum* f. sp. *lycopersici* (FOL).

The 133 shared orthogroups between the three FOP isolates, equates to 232 orthologous genes in FOP1 EMR, 226 in F81, and 184 in R2. In FOP1 EMR there were 79 putative effectors predicted by EffectorP in shared orthogroups with F81 and R2. In addition, there were five CAZymes and five genes within 2 kb of a mimp which were orthologous to F81 and R2 (Table A 3.1). In F81, there were 81 putative effectors predicted by EffectorP, three CAZymes and nine genes within 2 kb of a mimp in shared orthogroups with FOP1 EMR and R2 (Table A 3.2). Of the 184 genes in R2 in shared orthogroups, 63 were putative effectors predicted by EffectorP, four were identified as CAZymes and there were nine genes within 2 kb of a mimp (Table A 3.2).

Genes in shared orthogroups across FOP isolates were filtered to only show genes that were predicted as secreted (Table 3.4) as they are the most likely candidates for putative effectors important in pathogenicity in FOP. There were different numbers of predicted secreted genes in shared orthogroups across the three isolates, with FOP1 EMR containing 13 and both F81 and R2 containing 11. There were 10 genes (in eight orthogroups, 1123, 11965, 12380, 12381, 12565, 12821 and 12822) identified which contained genes with a secretion signal that were shared by all three FOP isolates. The most likely candidates for putative effectors important for pathogenicity in FOP were those genes also located within 2 kb of a mimp. However, in FOP1 EMR there was only one gene (g17046) within 2 kb of a mimp, and the corresponding genes sharing the same orthogroup were not predicted to be secreted in F81 or R2. Therefore, it is likely that the six genes (in orthogroups 1123, 12017, 12821 and 12822) located on LS regions and also predicted as putative effectors by EffectorP in all three isolates were the most likely candidates for FOP specific effectors. Two of the genes in F81 and R2 in shared orthogroup 1123 were also located within 2 kb of a mimp.

Table 3.4 Shared predicted secreted genes identified by orthology analysis in each of the three isolates of *Fusarium oxysporum* f. sp. *pisi* (FOP). Each isolate belonged to a different putative FOP race type (FOP1 EMR = race 1, F81 = race 2 and R2 = race 5). The contig and location of the contigs in the genome were included as Core and LS (lineage specific). Functional annotation of genes including whether it was predicted as secreted (Sec.), within 2 kb of a mimp, identified by EffectorP, which orthogroup the gene belonged to and the number of genes in that orthogroup (gene counts) were included in the ratio FOP1EMR:F81:R2

Isolate	Gene	Contig	Location	Sec.	2 kb of a mimp	EffP	Orthogroup	Orthogroup gene counts
FOP1 EMR	g17700	22	LS	Yes		Yes	1123	6:3:3
FOP1 EMR	g21974	68	LS	Yes		Yes	1123	6:3:3
FOP1 EMR	g17039	19	LS	Yes		Yes	1123	6:3:3
FOP1 EMR	g17046	19	LS	Yes	Yes	Yes	2436	2:1:1
FOP1 EMR	g16069	16	Core	Yes			11965	2:1:1
FOP1 EMR	g16256	16	Core	Yes			11965	2:1:1
FOP1 EMR	g18161	24	Core	Yes		Yes	12017	1:1:1
FOP1 EMR	g19102	28	LS	Yes		Yes	12017	2:1:1
FOP1 EMR	g45	1	Core	Yes		Yes	12380	2:2:3
FOP1 EMR	g46	1	Core	Yes		Yes	12381	1:1:1
FOP1 EMR	g8816	6	Core	Yes			12565	1:1:1
FOP1 EMR	g19442	30	LS	Yes		Yes	12821	1:1:1
FOP1 EMR	g19443	30	LS	Yes		Yes	12822	1:1:1
F81	g18880	35	LS	Yes	Yes	Yes	1123	6:3:3
F81	g18891	35	LS	Yes	Yes	Yes	1123	6:3:3
F81	g15746	17	LS	Yes		Yes	1123	6:3:3
F81	g19323	40	LS	Yes			11965	2:1:1
F81	g19265	40	LS	Yes		Yes	12017	2:1:1
F81	g14418	15	Core	Yes		Yes	12380	1:1:1
F81	g14417	15	Core	Yes		Yes	12381	1:1:1
F81	g11159	10	Core	Yes			12565	1:1:1
F81	g19210	39	LS	Yes			12659	1:1:1
F81	g19747	46	LS	Yes	Yes	Yes	12821	1:1:1
F81	g19746	46	LS	Yes		Yes	12822	1:1:1
R2	g17990	20	LS	Yes	Yes	Yes	1123	6:3:3
R2	g18001	20	LS	Yes	Yes	Yes	1123	6:3:3
R2	g17971	20	LS	Yes		Yes	1123	6:3:3
R2	g17893	20	LS	Yes			11965	2:1:1
R2	g17944	20	LS	Yes		Yes	12017	2:1:1
R2	g3956	3	Core	Yes		Yes	12380	1:1:1
R2	g3955	3	Core	Yes		Yes	12381	1:1:1
R2	g6934	4	Core	Yes			12565	1:1:1
R2	g19237	33	LS	Yes			12659	1:1:1
R2	g18184	21	LS	Yes	Yes	Yes	12821	1:1:1
R2	g18183	21	LS	Yes		Yes	12822	1:1:1

Genes that were unique (shared no orthology with other FOP isolates) were also examined. There were 161 genes unique to FOP1 EMR (race 1), 89 unique in F81 (race 2) and 23 unique to R2 (race 5). These genes were filtered to only display those which were also secreted; therefore, genes listed were FOP race specific genes predicted to produce secreted proteins. In FOP1 EMR there were 10 unique predicted secreted genes (Table 3.5), with five also predicted as putative effectors by EffectorP (g15152, g21075, g21410, g21926, g21927). Three of these genes (g21410, g21926, g21927) were also located within 2 kb of a mimp. All putative effectors predicted by EffectorP and those within 2 kb of a mimp were all located on LS regions of the genome (3.3.2) supporting the hypothesis that they are race specific effector genes potentially important in infection.

There were fewer race specific genes in F81 and R2 with five and three genes respectively (Table 3.5). None of these genes were predicted as putative effectors by EffectorP or located within 2 kb of a mimp. Two genes from F81 (g20210 and g19595) were identified as CAZymes (glycoside hydrolase family 12), located in the LS regions of the genome, and two other genes (g8752 and g18567) were annotated as nucleoside triphosphate hydrolyases and tyrosine protein kinases (Table 3.5). In R2, the three identified unique genes were annotated as Willebrand factor type A proteins and Ankyrin repeat domains, two of which were located on LS regions of the genome, which were more likely to be candidate race specific effectors.

Table 3.5 Unique predicted secreted genes identified by orthology analysis in three isolates of *Fusarium oxysporum* f. sp. *pisi* (FOP). Each isolate belonged to a different putative FOP race type (FOP1 EMR = race 1, F81 = race 2 and R2 = race 5). Functional annotation of genes including whether it was predicted as secreted, within 2 kb of a mimp, identified by EffectorP and which orthogroup it belongs to was included. InterProScan results of possible protein function was also included if available.

Isolate	Gene	Contig	Loc.	Sec.	2 kb of a mimp	EffP	Orthogroup	Annotation
FOP1 EMR	g16492	17	Core	Yes			14588	
FOP1 EMR	g15152	13	LS	Yes		Yes	13674	
FOP1 EMR	g21075	45	LS	Yes		Yes	13674	
FOP1 EMR	g17853	22	LS	Yes			13749	
FOP1 EMR	g17854	22	LS	Yes			13749	
FOP1 EMR	g21714	58	LS	Yes			13822	
FOP1 EMR	g20503	39	LS	Yes			13599	
FOP1 EMR	g21410	51	LS	Yes	Yes	Yes	12056	
FOP1 EMR	g21926	65	LS	Yes	Yes	Yes	12056	
FOP1 EMR	g21927	65	LS	Yes	Yes	Yes	12056	
F81	g20210	58	LS	Yes			12504	CAZY:GH12
F81	g19595	44	LS	Yes			12504	CAZY:GH12
F81	g6524	5	Core	Yes			13587	
F81	g8752	8	Core	Yes			14576	Nucleoside triphosphate hydrolase
F81	g18567	32	LS	Yes			13587	Tyrosine protein kinase ephrin
R2	g3947	3	Core	Yes			14048	Willebrand factor type A
R2	g18967	28	LS	Yes			14048	Willebrand factor type A
R2	g15874	13	LS	Yes			12595	Ankyrin repeat containing domain

3.4 Discussion

In this chapter, genomes of three FOP isolates (FOP1 EMR, F81 and R2) representing three putative races (1, 2 and 5 respectively) were sequenced, assembled and putative effector genes identified. Core and LS regions of each FOP genome were distinguished based on synteny analyses with the published FOL genome. Known *SIX* genes were identified in FOP in different complements across the putative races, with all of them (apart from one gene in F81) being identified on LS regions of the genomes of each isolate. This suggests there could be a race specific complement of *SIX* genes as eight were identified in the race 1 isolate FOP1 EMR (*SIX1*, *SIX6*, *SIX7*, *SIX9*, *SIX10*, *SIX11*, *SIX12* and *SIX14*), five in the race 2 isolate F81 (*SIX1*, *SIX6*, *SIX9*, *SIX13* and *SIX14*) and three in the putative race 5 isolate R2 (*SIX1*, *SIX6* and *SIX13*). In addition, novel putative effectors were found that were non-orthologous to FOC and FOL, which could potentially be important in the pathogenicity of FOP in pea. Furthermore, unique non-orthologous genes predicted to be secreted were identified in each race (1, 2 and 5), which could be important for cultivar specific pathogenicity of each race. To our knowledge this is the first genomic study of multiple FOP races for effector discovery and genome synteny compared to FOL.

The assembly statistics for the FOP genomes were comparable to other *F. oxysporum* f. spp. such as FOC and FOL, suggesting successful sequencing and assembly of reads (Armitage *et al.*, 2018; Ma *et al.*, 2010). The assemblies (contig number and sizes) were variable between the three isolates, with R2 containing the fewest contigs, suggesting a marginally better assembly than the others. In a number of cases, whole chromosomes were inferred from single contigs in each isolate, suggesting that all the genomes were assembled to a high standard. Assembly quality was also supported by the percentage of the genome which was repeatmasked, as these harder to assemble regions were resolved, therefore resulting in a more contiguous assembly. BUSCO compares predicted genes to known fungal genes to assess assembly quality and gene space (Simão *et al.*, 2015). The results of 98-99% suggests that core genes were identified and therefore the genomes were assembled well. Effective gene prediction was helped by two prediction programmes, BRAKER1 (uses RNAseq for training gene models, Hoff *et al.*, 2016) and CodingQuarry (uses assembled or aligned RNAseq transcripts to predict protein-coding gene sequences, Testa *et al.*, 2015) to ensure that very few genes were missed.

The difference between the total number of putative effectors predicted by EffectorP in the FOP genomes was greater than those also predicted to be secreted, which was only around one tenth of the total identified. This is a result of EffectorP possibly producing false positive results, as expected when not combined with those genes also encoding a secretion signal, although it is widely thought to provide a reliable approach for identifying effectors (Sperschneider *et al.*, 2016). This was also the case for CAZymes as only around half the total number were predicted as secreted, as carbohydrate active enzymes are involved in other carbohydrate related functions occurring within host cells, as well as being involved in pathogenicity. Of course, some effectors may not have a secretion signal, as for example, the *SIX12* gene in FOL does not encode a protein with a signal peptide for secretion, yet the corresponding protein was found in xylem sap from FOL infected tomato plants so it is clearly secreted (Schmidt *et al.*, 2013). Candidate FOP effectors were identified by detecting genes predicted as secreted located within 2 kb of a mimp (see 3.1 Introduction) and comparable to those reported for FOC isolate FUS2 (31) and FOL (22) (Armitage *et al.*, 2018; Ma *et al.*, 2010). Few predicted secreted genes within 2 kb of a mimp were present in non-pathogenic *F. oxysporum* isolates such as Fo47, as Armitage *et al.* (2018) identified only three, suggesting that these genes are related to pathogenicity. Understanding the expression of these putative effectors would provide further evidence as to whether they could be related to pathogenicity and was explored in Chapter 4.

FOP genomes were compared to FOL to assess synteny using orthologous genes in a 1:1 relationship, visualised using circos plots. Contigs in each isolate were grouped into 11 inferred core chromosomes, matching the 11 core chromosomes in *F. oxysporum*, confirmed in FOL (Ma *et al.*, 2010), FOC (Armitage *et al.*, 2018), f. sp. *radicis-cucumerinum* (van Dam *et al.*, 2017) and FOP (Williams *et al.*, 2016). The ungrouped contigs based on these synteny plots were identified as FOP lineage specific regions, which shared little synteny with the FOL genome, and are therefore likely to contain genes specific for FOP pathogenicity. When comparing the synteny of FOP1 EMR and FOL, it was observed that there were contigs where gene order crossed over between the isolates; for example, chromosome two and four in FOL with contigs three and two in FOP1 EMR. This suggests that FOP has had some chromosomal rearrangement in some of its core chromosomes, or that there was some level of misassembly of these contigs in

FOP. This is also seen to a slightly lesser extent in FOP isolates F81 and R2, possibly supporting core chromosomal rearrangement at some point in the evolution of this f. spp.

The location of predicted genes on individual contigs was also determined. The majority of putative effectors (EffectorP) and CAZymes were located on the core genome, with some core contigs containing high densities of these genes for their comparative size. These genes are likely to be generally related to pathogenicity of *F. oxysporum* and potentially other fungal pathogens as they are located in the core chromosomes, which are likely to have been vertically inherited from a common ancestor (Ma, 2014). CAZymes for example, are commonly identified across a range of pathogenic *Fusarium* species and other fungi, and are well known to be involved with plant cell wall degradation, a vital part of the infection process (Sista Kameshwar & Qin, 2018). Enrichment of EffectorP and CAZyme genes in core chromosomes was also observed in the FOC genome as well as non-pathogenic *F. oxysporum* isolates (Armitage *et al.*, 2018), suggesting that these genes may be involved in plant root colonisation in multiple f. spp., (Armitage *et al.*, 2018). In contrast, the majority of the mimp-related effector genes identified across FOP genomes, and all but one of the *SIX* genes, were identified in LS regions, suggesting that these regions contribute to host specific virulence of FOP towards pea, and have the potential to evolve rapidly which could lead to the evolution of new races. The EffectorP, CAZymes and *SIX* genes as well as other putative effectors identified particularly in the LS regions, (especially those within 2 kb of a mimp) are therefore likely candidates for important FOP specific effectors. This is supported by other studies where LS regions in both FOC and FOL also contained the highest proportion of putative host-specific effectors (Armitage *et al.*, 2018; Ma *et al.*, 2010).

Homologs of *SIX* genes were identified in all three FOP races with each race having a specific complement of *SIX* genes. Homologs of *SIX* genes have been found in many other *F. oxysporum* f. spp. with examples such as: *SIX3*, *SIX5*, *SIX7*, *SIX9*, *SIX10*, *SIX12* and *SIX14* in FOC; *SIX1*, *SIX7* and *SIX8* in f. sp. *cubense* and *SIX6*, *SIX9*, *SIX11* and *SIX13* in f. sp. *radicis-cucumerinum* (Taylor *et al.*, 2016; Meldrum *et al.*, 2012; van Dam *et al.*, 2017). In a study by Fraser-Smith *et al.* (2014), it was reported that only *SIX13* and *SIX14* had previously been identified in FOP while Taylor *et al.* (2016) identified *SIX* genes across three races of FOP with race 1 containing five *SIX* genes (*SIX7*, *SIX10*, *SIX11*, *SIX12* and *SIX14*), race 2 containing two (*SIX13* and *SIX14*) and race 5 containing

only one (*SIX13*) *SIX* gene. Furthermore, another recent genome study of a FOP race 5 isolate found four *SIX* genes: *SIX1*, *SIX9*, *SIX13* and *SIX14* (Williams *et al.*, 2016). However, in the current study race 1 was found to contain *SIX1*, *SIX6*, *SIX7*, *SIX9*, *SIX10*, *SIX11*, *SIX12* and *SIX14*, race 2 contained *SIX1*, *SIX6*, *SIX9*, *SIX13* and *SIX14* and race 5 contained only *SIX1*, *SIX6* and *SIX13*. This variation in FOP *SIX* gene complements between different studies and the findings reported here could be due to whether a genome sequencing or PCR approach was used, and in the latter case, which primers were utilised to amplify different *SIX* genes. *SIX* gene homologs associated with different f. spp. vary in sequence and therefore may not be amplified by *SIX* gene primers designed for another f. spp. The location of the majority of *SIX* genes in LS regions of the FOP genomes supports the hypothesis that these are likely to be FOP specific effectors. However, *SIX9* in F81 was located on a core chromosome, which could have been transferred there by transposon activity from the LS regions. In FOL, two copies of *SIX8* were identified in the telomeres of two core chromosomes which was also hypothesised to have been due to transposons or recombination events (Fraser-Smith *et al.*, 2014; Ma *et al.*, 2010). The different complements of *SIX* genes in FOP races may have evolved due to loss of genes to evade detection by the host (Chapter 4). This has been observed in FOL where the loss of *SIX4* from race 1 resulted in the emergence of race 2, which avoids detection by the host resistance gene *I1* (Houterman *et al.*, 2008).

The majority of the mimps identified in the FOP genomes were located on LS contigs across all three isolates. The large amount of repetitive DNA in the LS regions likely contributed to their assembly into short contigs, as it can be difficult to assemble repetitive regions of the genome, even with long read sequencing technologies such as MinION, especially if these regions are longer than the sequencing reads (Tyson *et al.*, 2018). Mimps have been found in other *Fusarium* genomes, but the relatively high abundance of mimps in the *F. oxysporum* species complex suggests that they originated in this species complex, and may have transferred into other *Fusarium* species by horizontal transfer of part of an accessory chromosome (van Dam & Rep, 2017).

There were relatively few shared orthogroups (51) between FOP1 EMR (race 1) and R2 (race 5), compared to those shared between FOP1 EMR and F81 (race 2) (278) and F81 and R2 (116), possibly due to the putative race 5 isolates being in a different lineage to races 1 and 2. Based on *TEF* phylogeny (Section 2.3.2) races 1 and 2 were found to be

closely related, with race 2 possibly evolving from race 1 (supported by a potential loss of *SIX* genes in race 2, discussed further in Section 4.4). However, race 5 was found to potentially belong to a separate lineage from races 1 and 2 (*TEF* phylogeny, Section 2.3.2), suggesting convergent evolution of FOP races and therefore potentially less orthologous genes between race 5 and the other races. R2 (race 5) shares 290 orthogroups with FOC, more than four times those shared between R2 and FOL, suggesting a closer common ancestry to FOC rather than FOL, supported by *TEF* phylogeny (Section 2.3.2) as isolates of FOC are found in the same clade as R2.

There was a total of ten genes (in eight orthogroups) with a signal peptide for secretion that were identified in orthogroups unique to FOP (when compared with FOC and FOL). Of these, eight genes (in six orthogroups) were also identified as putative effectors by EffectorP, and six of these genes (in four orthogroups 1123, 12017, 12821 and 12822) were located on LS regions in all of the FOP isolates. These genes are the most likely candidates for putative effectors that are specific for pathogenicity in FOP.

Unique genes containing a signal peptide for secretion were also identified within each FOP race, in order to predict putative race specific effectors. The three unique secreted genes identified in FOP1 EMR were all in the same orthogroup, suggesting that this group contains race specific effectors. They were also all found on LS regions and within 2 kb of a mimp suggesting potential duplication (two genes were on the same contig) events due to their proximity to transposable elements. In F81, the two CAZymes identified as unique genes predicted to be secreted belonged to the glycoside hydrolase 12 (GH12) family, a group possessing both endoglucanase and xyloglucanase activities, which are both involved in the breakdown of plant cell components such as cellulose and xyloglucan for cell wall rigidity (Yang *et al.*, 2017). Cell wall degrading enzymes have been shown to be important in pathogenicity of FOL through knockout studies resulting in reduced virulence (Jonkers *et al.*, 2009); therefore unique CAZymes could be an important adaption for cultivar specific races of FOP. There were only three unique secreted genes identified in R2, one of which was annotated as an Ankyrin repeat possibly linked to protein binding modules (Schmidt *et al.*, 2013).

The putative effector genes identified through this genome analysis both shared between FOP races or unique to each race are potential candidates for gene knockout studies to understand their role in pathogenicity. There were many genes shared between R2 (race

5) and F81 (race 2), possibly supporting the horizontal gene transfer from race 2 to 5 (as suggested by *TEF* phylogeny, Section 2.3.2), or an incorrect assignment of races. As FOP races were identified based solely on *TEF* gene phylogenetic analysis (Section 2.3.2) and the differences in *SIX* gene complements in this chapter, further investigation was carried out to confirm race identity using pea differential cultivar pathogenicity tests in Chapter 5.

Overall, the genome assembly and analysis of three putative races (1, 2 and 5) of FOP has revealed the presence and location in the genome of putative effector genes, including the well characterised *SIX* genes identified in other *F. oxysporum* f. spp. To our knowledge this is the first study to compare multiple FOP races to try and determine unique genomic features between races. This work will be used in conjunction with qPCR and RNAseq data in the following chapter, to understand the expression of putative effectors in FOP.

4. Effector gene analyses in FOP as a potential method for determining race type

4.1 Introduction

Successful fungal colonisation and infection of plants relies on the ability of fungi to modify their host in order to sequester the necessary nutrients for growth. One mechanism for this is through the secretion of virulence determinants known as effector proteins (Selin *et al.*, 2016). Effectors are defined as molecules that alter the host cells structure and function, which then serve in either facilitating infection or can trigger the host defence response (Selin *et al.*, 2016). Most plant pathogens utilise small, cysteine rich, secreted proteins to manipulate and suppress the host defence response, typically through interactions with host proteins (Schmidt *et al.*, 2013). A particularly well studied set of effector proteins in *Fusarium oxysporum*, strongly implicated in pathogenicity, are the Secreted in xylem (Six) proteins, which were identified in the xylem sap of tomato plants upon infection by *Fusarium oxysporum* f. sp. *lycopersici* (FOL) (Houterman *et al.*, 2007). Six proteins are coded for by *SIX* genes, of which fourteen have been identified so far in FOL which are generally found on lineage specific regions (LS) of the genome (Ma *et al.*, 2010; Schmidt *et al.*, 2013).

The first Six protein identified in FOL was Six1, which was found to be essential for full virulence. It also behaves as an avirulence factor by triggering host defences through interaction with the tomato *I-3* resistance gene (Rep *et al.*, 2004; Stergiopoulos & de Wit, 2009). Six4 has also been found to confer avirulence to FOL strains on tomato lines carrying the *I* or *I-1* resistance gene but is not required for full virulence of FOL on plants that lack these resistance genes. Six4 also functions to suppress *I-2* and *I-3* mediated resistance in tomato (Houterman *et al.*, 2008). Six3 is also required for full FOL virulence in a susceptible host and also triggers resistance in tomato plants carrying the resistance gene *I-2* (Houterman *et al.*, 2009). Six3 and Six5 interact to overcome *I-2* mediated cell death and are therefore both required for full virulence (Ma *et al.*, 2015). Knockout studies with *SIX6* resulted in reduced virulence and is therefore necessary for complete virulence in FOL (Gawehns *et al.*, 2014). *SIX8 – 14* were identified in FOL after analysis

of the FOL genome for novel effectors (Schmidt *et al.*, 2013), and have since been found in other *formae speciales* (f. spp.) (Fraser-Smith *et al.*, 2014). Analysis of the FOL genome in relation to *SIX1 – SIX7* revealed association with transposable elements (TE), as in all cases a miniature impala (mimp) was found upstream, and mFot5 was frequently found downstream of the *SIX* gene open reading frame (ORF) (Schmidt *et al.*, 2013). Mimps and mFots are types of Miniature Inverted-repeat TEs (MITEs) which are DNA transposons (van Dam *et al.*, 2016) and a high concentration of mimps were found on chromosome 14 in FOL (the 'pathogenicity' chromosome) where most of the *SIX* genes are concentrated (Schmidt *et al.*, 2013).

Since the discovery of *SIX* genes in FOL and their implication in pathogenicity, they have also been identified in many other f. spp. of *F. oxysporum* (Fraser-Smith *et al.*, 2014). *SIX* genes have been found in isolates of many *F. oxysporum* f. spp., including f. sp. *cubense* (banana), *fragariae* (strawberry), *vasinfectum* (cotton), *niveum* (watermelon), *cepae* (onion) and *betae* (cabbage), along with many others (Fraser-Smith *et al.*, 2014; Meldrum *et al.*, 2012; Sasaki *et al.*, 2015; Taylor *et al.*, 2016; Covey *et al.*, 2014; Chakrabarti *et al.*, 2011). *SIX13* and *SIX14* have also been found in *F. oxysporum* f. sp. *pisi* (FOP), and therefore could potentially contribute to its pathogenicity (Fraser-Smith *et al.*, 2014). Three isolates of FOP (FOP1, FOP2 and FOP5 from the present study) representing three different races were included in a study of *SIX* gene presence in f. sp. *cepae* and other selected f. spp., which found that there were up to six *SIX* genes present across the races tested, and that there was distinct variation in *SIX* gene complement across the races (Taylor *et al.*, 2016). Although *SIX* gene homologues to those characterised in FOL have been identified in many f. spp. via PCR or genome sequencing, and implicated in infection in expression analyses *in planta*, there are likely to be many other important secreted proteins during infection which are unique to different f. spp.

The presence/absence and sequence of *SIX* genes in FOL has also been used to distinguish between different *F. oxysporum* f. spp. and also sometimes races within f. spp. For instance, *SIX6* sequence differences have been shown to distinguish FOL from other f. spp. such as f. sp. *melonis*, f. sp. *vasinfectum*, f. sp. *cucumerinum* and f. sp. *niveum* (Lievens *et al.*, 2009; van Dam *et al.*, 2016). Lievens *et al.* (2009) showed that *SIX4* was only found in FOL race 1 and therefore subsequent races evolved to evade host detection through a loss of this gene. FOL race 3 also acquired mutations in *SIX3*, allowing it to

evade host detection and therefore can be used to distinguish this race from races 1 and 2. (Lievens *et al.*, 2009). Sequence differences in *SIX8* have also been used to distinguish races in *F. oxysporum* f. sp. *cubense* (wilt pathogen of banana causing Panama disease) allowing race 4 to be distinguished from races 1 and 2, as well as differences between tropical and sub-tropical race 4 isolates (Fraser-Smith *et al.*, 2014). The *SIX6* gene has also been identified in *F. oxysporum* f. sp. *niveum*, where it is only present in races 0 and 1 and not in the more virulent race 2 (Niu *et al.*, 2016). Although distinction of races is possible using *SIX* genes, most of these studies still relied on pathogenicity tests with plant differentials to confirm race identity.

Presence of *SIX* genes in *F. oxysporum* genomes does not provide information on expression or help elucidate function. One approach to determine their role in pathogenicity is to examine expression levels over time *in planta* following infection. In *F. oxysporum* f. sp. *cepae* (FOC), homologs of *SIX3*, *SIX5*, *SIX7*, *SIX9*, *SIX10*, *SIX12* and *SIX14* were identified in all highly pathogenic isolates from onion (Taylor *et al.*, 2016). All seven of these *SIX* genes were expressed *in planta* across a time course following inoculation of onion seedling roots with FOC *in vitro*. Expression was highly upregulated at the early stages of infection, with peak expression at 36 – 72 hpi (Taylor *et al.*, 2016). Studies of *SIX4* in an *F. oxysporum* isolate affecting *Arabidopsis* revealed gene expression during infection of plant roots, but no expression in culture (Thatcher *et al.*, 2012). *SIX6* has also been found to be expressed at the early stages of FOL infection on tomato seedlings, and it was concluded that the presence of living host cells was required as there was no expression in mycelium on synthetic media (Gawehns *et al.*, 2014). Similarly, *SIX6* was also found to be expressed from 3 dpi in watermelon roots infected with *F. oxysporum* f. sp. *niveum* (Niu *et al.*, 2016). *SIX1* expression in FOL also requires the presence of living tomato cells for upregulated expression after inoculation (van der Does *et al.*, 2008). Many gene expression studies have identified *SIX* genes being expressed *in planta*, however, gene knockout studies are required to determine their role in pathogenicity.

Expression of other putative effectors in *F. oxysporum*, following genome analysis (Chapter 3), can be determined through transcriptome analyses using qPCR and RNAseq techniques. Whereas qPCR determines the expression of single target genes, RNAseq can be used to sequence cDNA derived from the entire transcriptome (Han *et al.*, 2015).

Observing the expression of *F. oxysporum* genes during infection of a susceptible host enables important effector candidates to be determined. RNAseq analysis has already been used to study effector gene expression in multiple f. spp. including FOL (Ma *et al.*, 2010; Schmidt *et al.*, 2013), f. sp. *melonis* (Sebastiani *et al.*, 2017), f. sp. *cubense* (Guo *et al.*, 2014), f. sp. *cepae* (Armitage *et al.*, 2018) and the legume infected f. sp. *medicaginis* (Thatcher *et al.*, 2016). *SIX* genes previously identified in f. sp. *medicaginis* (*SIX1*, *SIX8*, *SIX9* and *SIX13*, Williams *et al.*, 2016), were also found to be highly expressed *in planta* (Thatcher *et al.*, 2016).

The main aim of this chapter was to confirm the race specific complement of *SIX* genes in FOP (Chapter 3) and understand the expression of these and other putative effectors using qPCR and RNAseq analysis.

The specific objectives were:

1. Screen *F. oxysporum* isolates collected from peas for the presence of *SIX* genes
2. Determine the expression of *SIX* genes during the early stages of infection in pea over time.
3. Examine FOP gene expression during infection of pea through transcriptome analysis.

4.2 Materials and Methods

4.2.1 *SIX* gene screening of *F. oxysporum* isolates

Preliminary whole genome sequencing was carried out using Illumina MiSeq sequencing for the three historic isolates FOP1 (race 1), FOP2 (race 2) and FOP5 (race 5) (Linfield, 1994), and assembled by Dr Andrew Armitage at NIAB-EMR. Firstly, adapter sequences were removed using *fastq-mcf* (Aronesty, 2013) and then *de novo* assembly was performed using Spades v.3.5.0 (Nurk *et al.*, 2013). Quast (Gurevich *et al.*, 2013) was used to summarise assembly statistics, and then RepeatMasker (Smit *et al.*, 2013-2015) and TransposonPSI (Haas, 2007) were used to identify repetitive regions. Additional MinION genomes of FOP1 EMR (race 1), F81 (race 2) and R2 (race 5) were assembled (Chapter 3) and also used for primer design. Unmasked assembled sequences were imported into Geneious (v. 6.1.5) and used to conduct BLAST (Boratyn *et al.*, 2013) searches for the 14 *SIX* genes previously identified in FOL (sequences obtained from NCBI). Sequences of positive hits (plus 500 bp up and downstream) were aligned in MEGA7 and used to check the utility of previously published primers for each of the 14 *SIX* genes (Lievens *et al.*, 2009; Taylor *et al.*, 2016). Where these primers did not match target sequences in FOP, new primers were designed outside the coding region of the gene using Primer3Plus (Untergasser *et al.*, 2007) and checked for self-hybridisation potential using Eurofins Oligo Analysis Tool (Eurofins, 2016). Primers were used to screen 41 of the UK *F. oxysporum* root rot isolates from pea (selected according to differences based on *TEF* phylogenetic analysis (Section 2.3.2) and pathogenicity (in the seed inoculation test, Section 2.3.3.1 and 5.3.2)) as well as the 29 FOP isolates obtained from the USA, Algeria and Czech Republic, previously race typed as races 1, 2 or 6 (Section 2.3.1) (Table 4.1). PCR reactions for each *SIX* gene were set up using REDTaq® ReadyMix® (Sigma-Aldrich) in 20 µL reaction volumes with 1 µL of DNA and 1 µL each of 10 µM primers with the thermocycling conditions as follows: one cycle for 2 min at 94˚C; 30 cycles of 45 s at 94˚C, 30 s at primer annealing temperature (Table A 4.1) and 1 min at 72˚C, followed by one cycle of 7 min at 72˚C. PCR products were visualised using gel electrophoresis (1% agarose gel containing GelRed™ at 2 µL per 100 mL of gel).

4.2.2 *SIX* gene expression analysis using qPCR from pea roots inoculated with FOP

4.2.2.1 Inoculation of pea roots and sampling over time

An *in vitro* method where peas were grown and inoculated with FOP isolates on an agar medium was adapted from Taylor *et al.* (2016). Autoclaved ATS medium (1M KNO_3, 1M KPO_4, 1M $MgSO_4$, 1M $Ca(NO_3)_2$, 20mM Fe-EDTA, 70mM H_3BO_3, 14mM $MnCl_2$, 0.5mM $CuSO_4$, 1mM $ZnSO_4$, 0.2mM Na_2MoO_4, 10mM NaCl, 0.01mM $CoCl_2$, 0.45% Gelrite (Duchefa Biochemie, Haarlem, The Netherlands) was used to three-quarter fill square petri dishes (12 x 12 x 1.7 cm, Greiner Bio-One, UK) and once set, the top 4.5cm of the gel was removed with a sterile spatula. Pea seeds (cv. Little Marvel) were sterilised in a 10% bleach (4.5 g 100 mL^{-1} sodium hypochlorite, 1 – 5% available chlorine)/sterile water (v/v) solution for 5 min, then rinsed with sterile distilled water (SDW) until no bubbles remained. Six seeds were pushed into the cut edge of the agar at approximately 1.5 cm intervals across the plate. Plates were sealed with tape, then wrapped in cling film in stacks of 6-8 and incubated at 20˚C for 3 days in the dark, then a further 5 days in light / dark (16 h day length). Spore suspensions of FOP1 EMR (race 1), F81 (race 2) and R2 (race 5) were made as described for the pathogenicity tests (Section 2.2.3.1), by releasing spores from 2-week-old agar plates with 10 mL SDW and filtering through three layers of miracloth. Spore suspensions were adjusted to 1 x 10^6 spores mL^{-1} using SDW with the addition of 200 µL of Tween20 L^{-1} and 1 mL pipetted directly onto pea roots and spread by tilting the plate. Plates were re-sealed and returned to the 20˚C (16 h photoperiod) incubator in the arrangement of a split pot design. In the split plot design individual plates were arranged in sealed packs (with cling film) containing one plate of each isolate, and one pack of these was situated on each of the four shelves in the incubator. One pack was removed per shelf at each sampling time to ensure any variation across shelves was accounted for. Pea root samples were taken at 10 time points (4, 8, 16, 24, 36, 48, 72 and 96 hours post inoculation), including a 0 h time point (pre-inoculation) and an uninoculated control (SDW/Tween only) collected at 96 hpi. Four plates were inoculated for each FOP isolate for collection at each timepoint when the roots of four plants (most established) per plate were removed, rinsed in SDW, flash frozen in liquid nitrogen and stored at -80˚C until use.

4.2.2.2 RNA extraction and cDNA synthesis

Pea roots from 4.2.2.1 were ground to a fine powder using a pestle and mortar filled with liquid nitrogen and approximately 100 mg of tissue transferred to a 2 mL tube. Frozen root material was ground further using a Dremel drill (model 398, with a rounded drill bit) and then RNA extracted using Trizol ® reagent (Ambion, Thermo Fisher Scientific) following the manufacturer's guidelines. Extracted RNA was precipitated using 900 µL of lithium chloride to 100 µL RNA (250 µL $LiCl_2$ + 650 µL DEPC treated water) to remove contaminants. Remaining DNA was removed from samples using DNase 1 (Sigma-Aldrich). RNA samples were visualised on a 2% agarose gel (containing GelRed™ at 2 µL per 100 mL of gel) with the addition of loading dye (Orange G, Sigma-Aldrich), to check for degradation. First strand cDNA was synthesised using Superscript II reverse transcriptase (Invitrogen, Thermo Fisher Scientific) following the manufacturer's protocol.

4.2.2.3 Quantitative PCR analysis of *SIX* gene expression

The expression of *SIX* genes identified in FOP was assessed using quantitative PCR of the cDNA derived from each of the inoculated pea root samples at each time point from 4.2.2.2. Published primers (Taylor *et al.*, 2016) were used for *SIX7*, *SIX9*, *SIX10*, *SIX12* and *SIX13* and the remainder (*SIX1*, *SIX6*, *SIX11* and *SIX14*) were designed in this study as described in 4.2.1 above, or by manually selecting candidate primers. Self-hybridisation potential was tested as well as any ability for DNA secondary structures to form (Zuker, 2003). Due to the multi-copy nature of *SIX6* in FOP race 1 (FOP1 and FOP1 EMR), primers were specifically designed to only amplify the single copy found in the other isolates FOP2/F81 (race 2) and FOP5/R2 (race 5). Additionally, due to multiple copies of *SIX1* being found, only one copy found in both F81 (race 2) and R2 (race 5) was used for primer design, as it was not possible to design unique primers for the copy in FOP1 EMR due to the similarity of sequences. Reverse transcription qPCR was performed in a Roche Lightcycler 480 using the Lightcycler 480 SYBR Green 1 Master mix (Roche, UK), following the manufacturer's protocol. All primers were used at a final concentration of 0.4 µM (except qSIX11 and qSIX14 which were used at 0.2 µM and qSIX6 which was used at 0.15 µM), using the following conditions: one cycle of 95˚C

for 5 mins, 45 cycles of 95˚C for 10 s, primer annealing temp (Table A 4.1) for 10 s and finally 72˚C for 10 s. The melt curve analysis follows this amplification stage to confirm the presence of a single PCR product. Standard curves were plotted for each gene, by using serially diluted genomic DNA, and the concentration of each gene expressed relative to the same housekeeping gene (*TEF*) as used for isolate identification (Chapter 2). In addition *TEF* has been shown to be stably expressed (Lysøe *et al.*, 2008) and has been used in previous similar studies (Taylor *et al.*, 2016).

4.2.2.4 Statistical analysis

All statistical analyses were carried out in Genstat® (release 18.1, VSN international Ltd). qPCR was used to determine the concentration of DNA for different *SIX* genes at different time points during infection of FOP in pea roots. The concentration of DNA for each *SIX* gene and the housekeeping gene *TEF* was calculated from the standard curve as above. DNA concentration values for the *SIX* genes were averaged across replicates and divided by the averaged DNA concentration for *TEF* for corresponding samples. These relative concentration values were then $\log_e$ transformed to account for increased variance across the time course. Analysis of variance (ANOVA) was then carried out using the transformed relative DNA concentration for each *SIX* gene, considering the blocking structure in the incubator. The split plot blocking structure (Section 4.2.2.1) was applied to the ANOVA when there were two or more isolates included in the analysis (for example, a *SIX* gene was present in two isolates, so the analysis included more than one treatment). When there was only one isolate included in the analysis (when a gene was unique to an isolate, e.g. *SIX7* in FOP1 EMR) a simpler randomised block design was used for the blocking structure of the incubator in the ANOVA. The overall effect of sampling time and isolate were observed for the expression of each gene. Significant differences between individual time points or isolates were compared with the least significant difference (LSD) at the 5% level. Back transformed means were calculated and plotted on graphs.

4.2.3 RNAseq expression analysis of FOP inoculated pea roots at 96 hours post inoculation

RNA extracted from pea roots from each replicate of the 96 hpi time point for FOP1 EMR (race 1), F81 (race 2) and R2 (race 5) (Section 4.2.2.2) was used to generate RNAseq data in order to assess overall FOP gene expression at the early stages of infection *in planta*. In addition, RNA was extracted from 3 replicate plates of mycelium of each isolate grown for 2 weeks on potato dextrose agar (PDA) plates as a control (methods as in Section 4.2.2.2) and treated with lithium chloride and DNase 1 (Sigma-Aldrich) as above (Section 4.2.2.2) before using for RNAseq library preparation. RNA integrity was determined using an Agilent Bioanalyser and library preparation was carried out using the Illumina TruSeq RNA V2 kit with the starting amount of total RNA normalised to 500 ng (Genomics facility, University of Warwick). RNA sequencing was carried out by Oxford Genomics Centre using an Illumina HiSeq4000 machine using 150 bp paired-end runs, with libraries pooled and ran over three lanes to give approximately 25 million pairs of reads (mycelium) and 50 million pairs of reads (pea roots) per sample.

All RNAseq analysis was performed in R apart from the trimming and alignment stages which were carried out on the NIAB-EMR servers, with the aid of Dr Andrew Armitage. Raw RNAseq reads were trimmed using *fastq-mcf* (Aronesty, 2013) and aligned to gene models (Chapter 3) using Salmon (Patro *et al.*, 2017) to quantify gene expression levels using FPKM (fragments per kilobase million). Sample-to-sample distances (similarities and differences between samples) were calculated using rlog transformed count data and visualised as a heatmap. In addition, a principal component analysis (PCA) was conducted in order to compare similarities and differences in sample transcripts, and plotted in terms of PC1 and PC2 showing the highest and second highest amounts of variation within samples respectively. DESeq2 analysis (Anders & Huber, 2010; Huber *et al.*, 2015) was performed using the 96 hpi pea root samples and mycelium from each FOP isolate and FPKM values in order to identify up and downregulated genes which showed a greater than two-fold change in expression. Mean FPKM values were taken for each gene from the three replicates for the 96 hpi pea root sample and mycelium (from plates) data for each isolate. FPKM values for upregulated and downregulated genes were arranged in descending order to identify the 30 highest expressed up/down regulated genes, which were then annotated with any functional analyses (Chapter 3). In addition,

the expression of *SIX* genes was examined for each FOP isolate and cross referenced to annotations from the genome analyses (Chapter 3). Genome annotations were also applied to identify any predicted genes which were secreted and found within 2 kb of a mimp, as these have been previously used to identify putative effectors in *F. oxysporum* (Schmidt *et al.*, 2013).

4.3 Results

4.3.1 *SIX* gene screening of *F. oxysporum* isolates

A selection of isolates from infected UK pea fields and previously race typed isolates from the USA, Algeria and Czech Republic were screened for the presence of *SIX* genes to determine whether there was a race specific complement between isolates.

Conducting BLAST searches of the assembled MiSeq genomes of isolates FOP1, 2 and 5, and the additional MinION assembled genomes of FOP1 EMR (race 1), F81 (race 2) and R2 (race 5) revealed differences in *SIX* gene profiles between the six isolates. FOP1/FOP1 EMR (race 1) was found to contain eight *SIX* genes (*SIX1*, *6*, *7*, *9*, *10*, *11*, *12* and *14*), with an additional copy of *SIX6* (*SIX6*[1]), unique to this isolate. FOP2 and F81 (race 2) had positive hits for five *SIX* genes (*SIX1*, *SIX6*[2], *SIX9*, *SIX13* and *SIX14*), with an additional copy of *SIX1*. FOP5 and R2 (race 5) had a similar profile as FOP2/F81 but lacked *SIX14* and had two additional copies of *SIX1* (instead of one additional copy). The above sequences were used to design specific primers and used in conjunction with previously published primers (Table A 4.1) to screen selected isolates.

PCR screens of the 29 previously race typed FOP isolates (Section 2.3.1) revealed that 27 contained at least one *SIX* gene, with 21 of these having the same *SIX* gene profile to one of the three genome sequenced isolates above. Of these, five isolates had the same complement of *SIX* genes as found in FOP1/FOP1 EMR and were therefore designated as preliminary FOP race 1 isolates (Table 4.1). A further 11 isolates had the same complement of *SIX* genes as FOP2/F81 and were assigned as preliminary FOP race 2 isolates, with the remaining four isolates having the same *SIX* gene profile as FOP5/R2, therefore being designated as preliminary FOP race 5 (Table 4.1). The final six isolates of the 27 overseas isolates only contained either *SIX6*[2] or *SIX14*, which did not match the complement of genes identified in any of the genome-sequenced FOP races. A similar result was observed for the 41 isolates from diseased UK peas, with only 16 isolates containing either *SIX6* or *SIX14*, which again did not match the profile of the designated race typed FOP isolates (Table 4.1). *F. oxysporum* isolates containing only one or two *SIX* genes were therefore designated as putative root rot isolates, as they had no similarity to any of the FOP isolates.

Table 4.1 Presence/absence of *Secreted In Xylem* (*SIX*) genes found in *Fusarium oxysporum* isolates from diseased UK pea samples and in *F. oxysporum* f. sp. *pisi* (FOP) isolates from overseas (with race designation) as determined by PCR. Isolates with a preliminary race type assigned before arrival (based on pea differential tests) are re-ordered by race based on *SIX* gene profiles. ■ = band (gene) present, ■ = faint band present, empty box = no band present.

Race*	Isolate	Location	Clade (*TEF*)	*SIX* gene presence/absence										Race (*SIX*)
				1	6[1]	6[2]	7	9	10	11	12	13	14	
1	FOP1	UK	Clade 7, containing potential race 1 isolates	■	■	■	■	■	■	■	■		■	1
1	FOP1 EMR	UK		■	■	■	■	■	■	■	■		■	1
1	F79	USA		■	■	■	■	■	■	■	■		■	1
1	CBS170.30	USA		■	■	■	■	■	■	■	■		■	1
1/2	Fw-09-C	USA		■	■	■	■	■	■	■	■		■	1
1/2	Fw-09-D	USA		■	■	■	■	■	■	■	■		■	1
2	FOP2	UK	Clade 8, containing potential race 2 isolates	■		■		■				■	■	2
2	F81	USA		■		■		■				■	■	2
2	F231	USA		■		■		■				■	■	2
2	F31	USA		■		■		■				■	■	2
2	F234	USA		■		■		■				■	■	2
2	F235	USA		■		■		■				■	■	2
2	F16	USA		■		■		■				■	■	2
2	F232	USA		■		■		■				■	■	2
2	F236	USA		■		■		■				■	■	2
2	F35	USA		■		■		■				■	■	2
2	F30	USA		■		■		■				■	■	2
2	F237	USA		■		■		■				■	■	2
2	F233	USA		■		■		■				■	■	2
5	FOP5	UK	Clade 4, containing potential race 5 isolates	■		■						■		5
2	R2	CR		■		■						■		5
2	F40	USA		■		■						■		5
2	PDA3b	USA		■		■						■		5
2	F42a	USA		■		■						■		5
2	Fw-08-04	USA	Clade 6	■		■								RR
1	CBS183.35	UN**	Clade 4	■		■							■	RR
1	AC1.2	ALG	Clade 5			■							■	RR
6	AC6	ALG	Clade 5											RR
2	CBS260.57	UN**	Clade 5											RR
2	Fw-08-02	USA	Clade 6	■		■								RR
2	Fw-08-03	USA	Clade 6	■		■								RR
1/2	Fw-09-E	USA	Clade 7			■		■					■	RR
	PG1	UK Field	Clade 1											RR
	PG2	UK Field	Clade 4											RR
	PG3	UK Field	Clade 1			■								RR
	PG4	UK Field	Clade 2											RR
	PG15	UK, CC	Clade 3			■								RR

Race*	Isolate	Location	Clade (*TEF*)	*SIX* gene presence/absence										Race (*SIX*)
				1	6[1]	6[2]	7	9	10	11	12	13	14	
	PG16	UK, CC	Clade 1											RR
	PG18	UK, CC	Clade 1											RR
	PG19	UK, CC	Clade 3			■								RR
	PG21	UK, CC	Clade 1											RR
	PG57	UK Field	Clade 1											RR
	PG58	UK Field	Clade 2										■	RR
	PG59	UK Field	Clade 2										■	RR
	PG60	UK Field	Clade 2											RR
	PG61	UK Field	Clade 2										■	RR
	PG62	UK Field	Clade 1											RR
	PG63	UK Field	Clade 5											RR
	PG65	UK Field	Clade 1											RR
	PG67	UK Field	Clade 3		■									RR
	PG72	UK Field	Clade 1											RR
	PG73	UK Field	Clade 1											RR
	PG74	UK Field	Clade 2											RR
	PG76	UK Field	Clade 2										■	RR
	PG77	UK Field	Clade 1											RR
	PG79	UK Trial	Clade 3			■								RR
	PG85	UK Field	Clade 2											RR
	PG108	UK Field	Clade 6											RR
	PG110	UK Field	Clade 1											RR
	PG113	UK Field	Clade 5											RR
	PG242	UK Field	Clade 1											RR
	PG247	UK Field	Clade 6										■	RR
	PG301	UK Field	Clade 1			■								RR
	PG316	UK Field	Clade 6											RR
	PG327	UK Field	Clade 3			■								RR
	PG336	UK Field	Clade 1			■								RR
	PG337	UK Field	Clade 1			■								RR
	PG389	UK Field	Clade 3			■								RR
	PG467	UK Field	Clade 1											RR
	PG476	UK Field	Clade 1											RR
	PG480	UK Field	Clade 1											RR
	PG494	UK Field	Clade 1											RR

SIX6[1] copy 1 (found only in FOP1/FOP1 EMR, and not FOP2/F81 or FOP5/R2)
SIX6[2] copy 2 (found in all genomes)
*Race of isolates as previously assigned before arrival. Empty boxes designate isolates not previously race typed.
**UN (Unknown location)
CR (Czech Republic)
ALG (Algeria)
UK CC (crop clinic PGRO)
UK Trial (Collected from pea plants grown in a field trial, Peterborough)
UK Field (Collected from peas from commercial fields across the UK)
RR (Root rot isolates with no FOP race classification)
Clade (*TEF*) based on clustering of isolates into clades using *TEF* gene sequencing (Figure 2.4)

4.3.2 *SIX* gene expression analysis using qPCR from pea roots inoculated with FOP

The three pathogenic genome sequenced FOP isolates (FOP1 EMR, F81 and R2; Chapter 3), representing the preliminary races 1, 2 and 5 respectively, were used to inoculate pea roots, and *SIX* gene expression examined over time following RNA extraction, cDNA synthesis and reverse transcription qPCR.

Visual symptoms of infection were present on pea roots from 24 hpi (across isolates tested), in the form of very mild browning, which increased in colour intensity and coverage of the root over time, until nearly all of the root was affected at 96 hpi (Figure 4.1). (F81 Figure A 4.1, R2 Figure A 4.2).

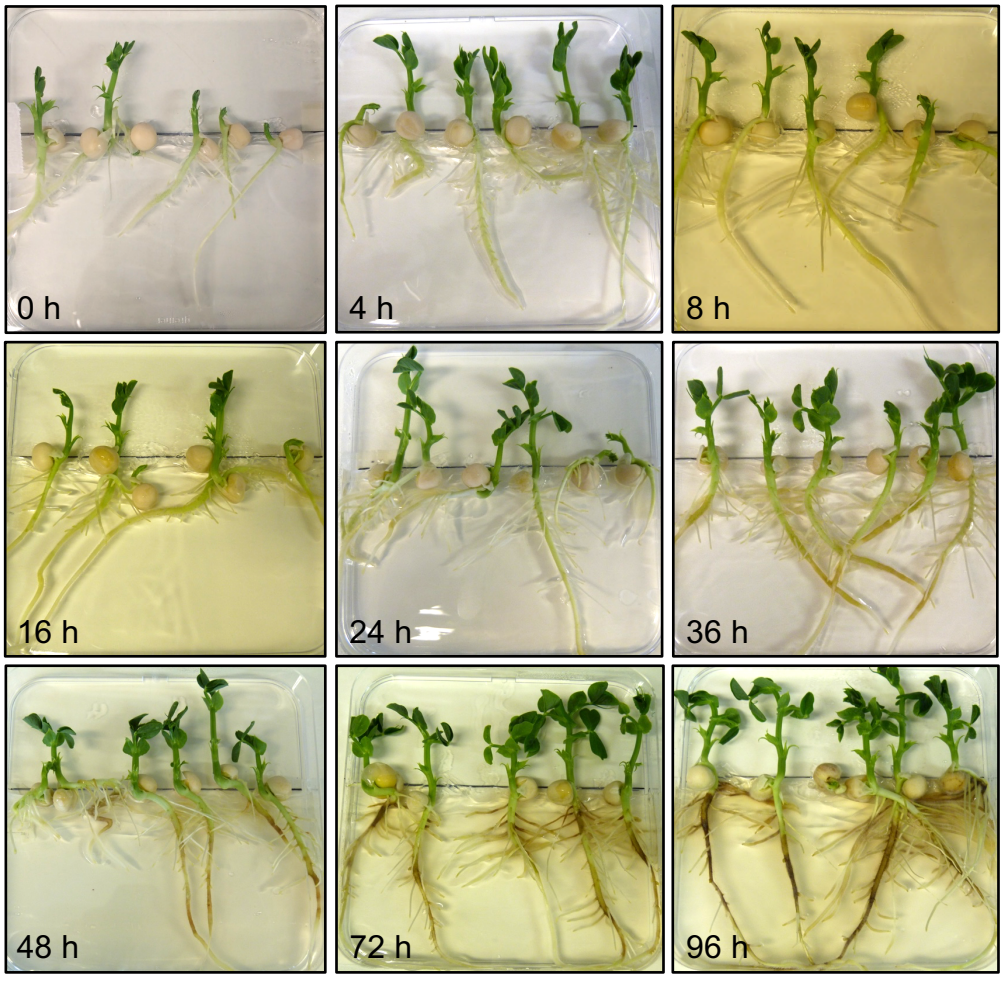


Figure 4.1 Symptoms of root browning over time on pea seedlings grown in square petri dishes and inoculated with *Fusarium oxysporum* f. sp. *pisi* (FOP) race 1 isolate FOP1 EMR.

For all of the *SIX* genes tested (apart from *SIX13*), there were overall significant differences in relative expression ($p < 0.001$; $p = 0.003$ for *SIX1*) over the time course using ANOVA. *SIX7*, *SIX10*, and *SIX12* expression (present in FOP1 EMR only) was first detectable at 24 hpi, and then increased at every time point over the experiment, with significant increases observed from 72 hpi onwards (Figure 4.2; Table 4.2). However, expression of *SIX6* for FOP1 EMR (race 1) was first detected at 36 hpi, which was in contrast to F81 (race 2) and R2 (race 5) where first detection was at 24h and 16h respectively (Figure 4.2; Table 4.2). Expression levels of *SIX6* for FOP1 EMR were also significantly less than for F81 and R2 over the subsequent time points (Figure 4.2; Table 4.2).

Relative expression of *SIX11*, present in FOP1 EMR, only was first detectable at 8 hpi, and significantly increased at every time point (with the exception of 36 to 48 hpi). *SIX13* (present in F81 and R2 only) expression in F81 was much greater at 4 hpi compared to the other *SIX* genes, but then decreased significantly to similar levels as seen in isolate R2. *SIX14*, present in FOP1 EMR and F81, was detectable from 16 hpi in F81 and relative expression increased over time until 48 hpi, followed by a plateau of expression until 96 hpi. However, for FOP1 EMR too few replicates had expression levels above the limit of detection (data not shown). Similarly, the expression of *SIX9* in FOP1 EMR was not detectable, however it was detectable 36 hpi in F81, where there were significant increases in expression levels at all time points other than 72 - 96 hpi, where expression plateaued.

Due to difficulty in designing primers for *SIX1* (Section 4.2.2), only F81 and R2 were included in the analysis for this gene, although *SIX1* was present in all three isolates, the sequence of *SIX1* in FOP1 EMR varied from the copies in F81 and R2 (4.3.1). Expression of *SIX1* was detected at 8 hpi for both isolates, with levels peaking at 48 hpi for F81 and at 72 hpi for R2. The expression levels for *SIX1* in F81 were significantly lower than in R2 at 36 hpi, 72 hpi and 96 hpi, comparing 5% LSD^2 (Table 4.2).

All *SIX* gene qPCR reactions also included a control (roots inoculated with SDW + Tween collected at 96 hpi) but no *SIX* gene expression was seen in any of these samples (data not shown).

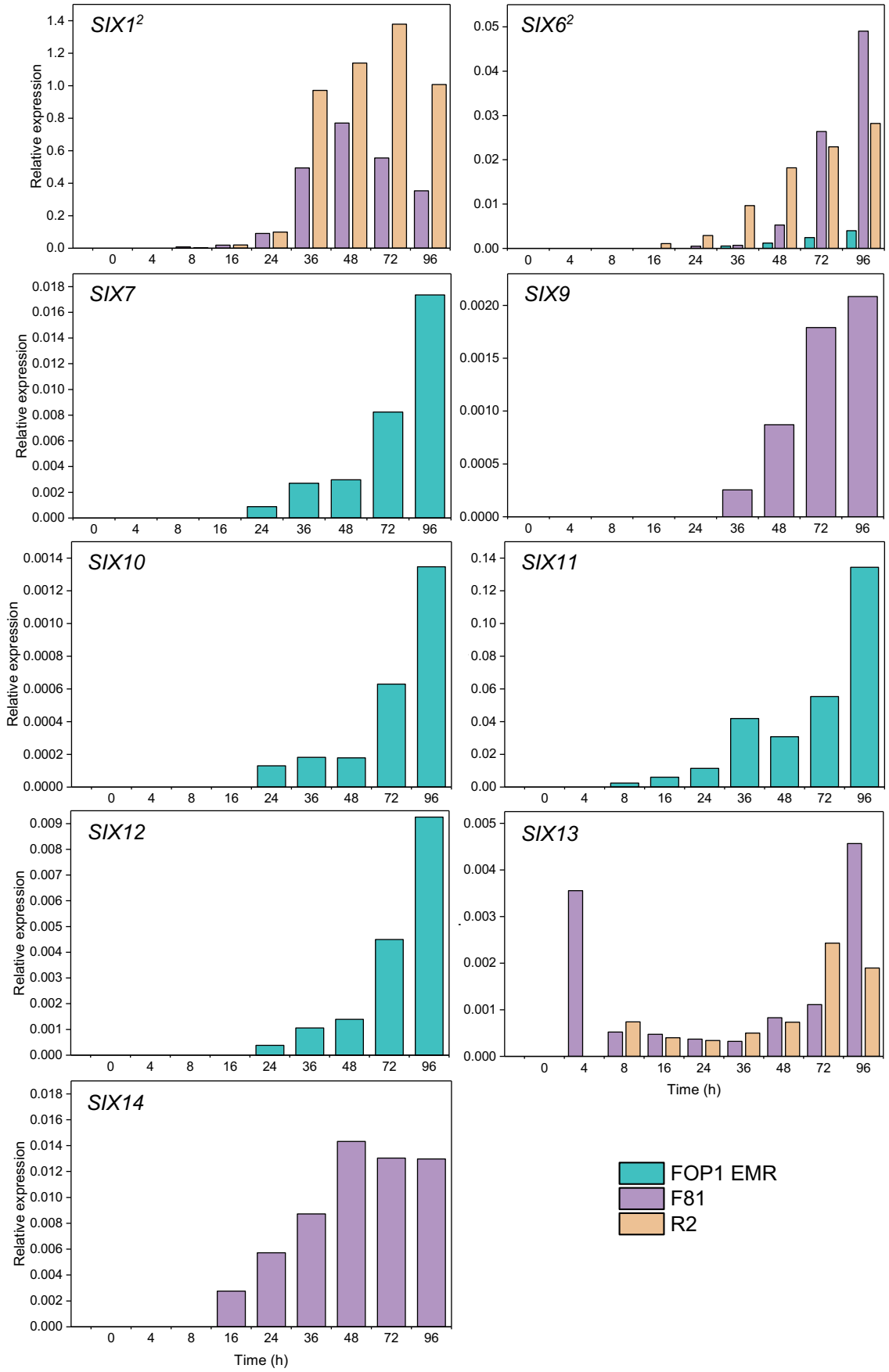


Figure 4.2 Quantitative expression of *Secreted In Xylem* (*SIX*) genes for RNA extracted from pea roots infected with *Fusarium oxysporum* f. sp. *pisi* (FOP) isolates FOP1 EMR (race 1), F81 (race 2) and R2 (race 5) as determined by reverse transcription qPCR. Expression was calculated relative to the *translation elongation factor 1α* (*TEF*), and $\log_e$ transformed before ANOVA. Values plotted represent the back transformed means of relative gene expression at nine time points (0 – 96 h) post inoculation.

Table 4.2 Log_e transformed ANOVA means of the expression of *Secreted In Xylem* (*SIX*) genes relative to the *translation elongation factor 1α* (*TEF*) gene for RNA extracted from pea roots infected with *Fusarium oxysporum* f. sp. *pisi* (FOP) isolates FOP1 EMR (labelled F1) (race 1), F81 (race 2) and R2 (race 5) between 0 – 96 hpi. Significant differences between time points can be calculated using the 5% LSD[1], and between isolates for the same time point using the 5% LSD[2] (*SIX1*, *SIX6* and *SIX13* only).

Time (h)	Relative expression of *SIX* gene to *TEF* (transformed means)												
	1			**6**		**7**	**9**	**10**	**11**	**12**	**13**		**14**
	F81	R2	F1	F81	R2	F1	F81	F1	F1	F1	F81	R2	F81
0	-	-	-	-	-	-	-	-	-	-	-	-	-
4	~	~	~	-	~	~	-	-	~	~	-5.64	~	~
8	-4.81	-5.67	-	~	~	~	-	~	-6.01	~	-7.56	-7.21	~
16	-3.94	-3.93	~	~	-6.83	~	~	~	-5.12	~	-7.66	-7.83	-5.89
24	-2.40	-2.31	=	-7.60	-5.85	-7.03	~	-8.95	-4.47	-7.88	-7.90	-7.98	-5.16
36	-0.71	-0.03	-7.52	-7.26	-4.64	-5.91	-8.27	-8.61	-3.17	-6.85	-8.04	-7.60	-4.75
48	-0.26	0.13	-6.72	-5.24	-4.01	-5.82	-7.05	-8.63	-3.48	-6.57	-7.10	-7.22	-4.25
72	-0.59	0.32	-6.01	-3.64	-3.78	-4.80	-6.32	-7.37	-2.89	-5.41	-6.80	-6.02	-4.34
96	-1.04	0.01	-5.53	-3.02	-3.57	-4.05	-6.17	-6.61	-2.01	-4.68	-5.39	-6.27	-4.35
LSD^1	0.73			0.56		0.62	0.70	0.72	0.36	0.45	0.70		0.45
LSD^2	0.61			0.42							0.70		

- Denotes missing values
~ Denotes average values containing missing values and those lower than the limit of detection of the primer pair
= Denotes all values below limit of detection for each primer pair

4.3.3 RNAseq expression analysis of FOP inoculated pea roots at 96 hours post inoculation

4.3.3.1 Assessing sample quality

In addition to examining relative *SIX* gene expression with quantitative PCR, an RNAseq approach was also taken to examine expression levels of these and other genes that may also be associated with FOP infection. RNAseq was therefore carried out for the same three isolates (FOP1 EMR, F81 and R2) using the infected pea root samples at 96 hpi, and on mycelium grown on PDA for approximately 2 weeks.

To assess overall similarity of samples, the distance matrix between samples was calculated and plotted as a heatmap. Distances between the 96 hpi time point root samples and the mycelium samples indicated that replicates for each isolate sample were highly similar. However, there were substantial differences in gene expression levels between inoculated pea samples and mycelium samples (from agar cultures) which showed no similarity with each other (Figure 4.3).

To further examine sample distances, a principal component analysis was carried out. There were large variances for principal component 1 (PC1), that indicated that there were considerable differences between the inoculated pea samples (96 h time point) and mycelium for each FOP isolate (Figure 4.4). As expected, much smaller differences were evident between replicates, with close clustering in 96 hpi FOP1 EMR samples, and very small percentage variance for F81 (2%) and R2 (1%) between replicates in the 96 hpi samples. One mycelium replicate (FOP1 EMR 3) did not cluster closely with the other replicates, indicating there was some variation between replicates.

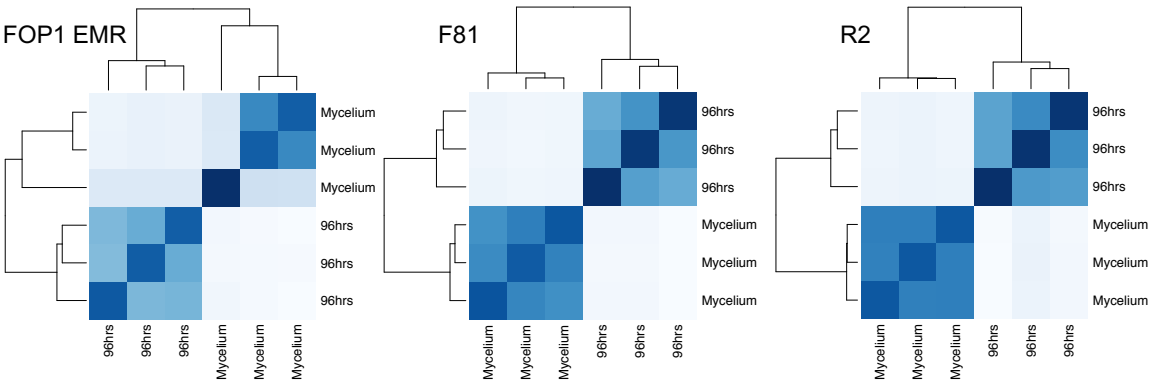


Figure 4.3 Sample distances heatmap of normalised RNAseq reads from infected pea roots (96 hpi) and agar-grown mycelium samples for *Fusarium oxysporum* f. sp. *pisi* (FOP) isolates FOP1 EMR, F81 and R2 isolates aligned to their respective genomes.

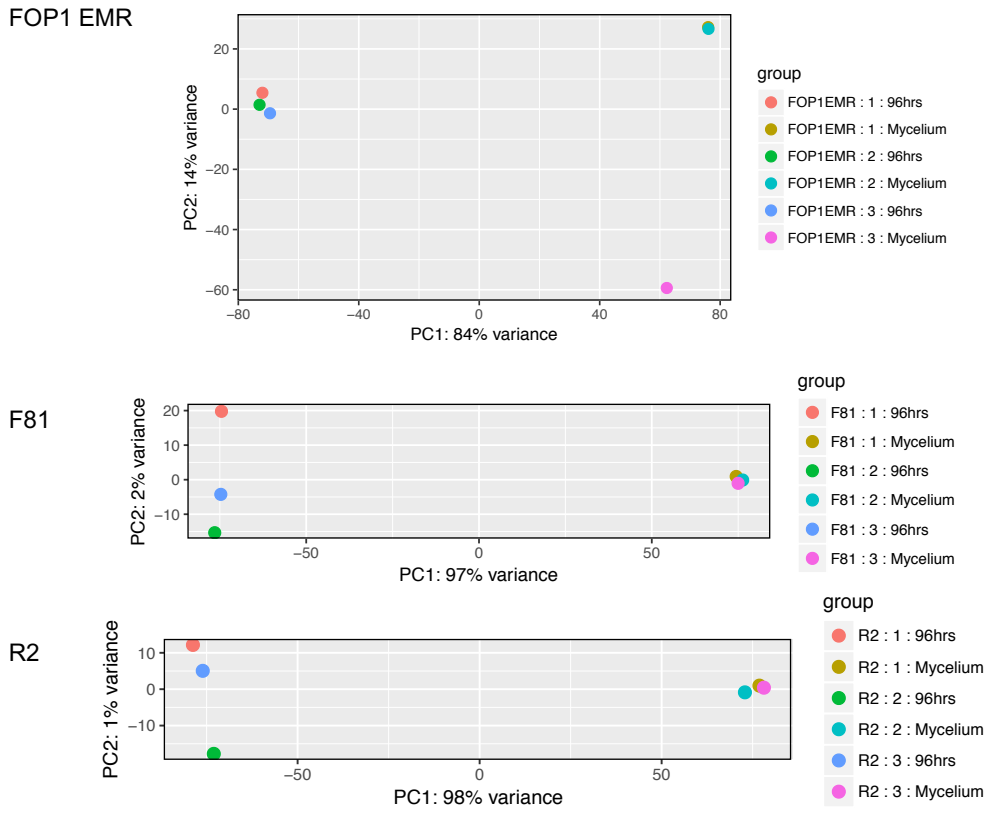


Figure 4.4 Principal component analysis (PCA) of normalised RNAseq reads from infected pea roots (96 hpi) and agar-grown mycelium samples for *Fusarium oxysporum* f. sp. *pisi* (FOP) isolates FOP1 EMR, F81 and R2 isolates aligned to their respective genomes.

4.3.3.2 Identification of differentially expressed genes in FOP

DESeq2 was used to identify significantly differentially expressed genes ($p < 0.05$) between 96 hpi pea root samples and agar-grown mycelium samples. Subsequently, average gene expression was calculated for the 96 hpi and mycelium samples separately (for genes considered differently expressed), and genes with more than an average 2-fold change in expression (between 96 h and mycelium) were considered as being up or down regulated. For FOP1 EMR, 7974 genes were differentially expressed with 3492 upregulated and 4455 genes being downregulated. For F81, there were 8130 differentially expressed genes at 96 hpi, with 3871 of these being upregulated and the remaining 4259 being downregulated. R2 showed the greatest number of differentially expressed genes at 8471, with 3971 of these being upregulated during infection (Table 4.3).

Table 4.3 Differentially expressed genes from extracted RNA for three *Fusarium oxysporum* f. sp. *pisi* (FOP) isolates, FOP1 EMR, F81 and R2 (race 1, 2 and 5 respectively), used to inoculate pea roots, and sampled 96 hpi. Upregulated and downregulated genes were calculated based on two-fold change from mycelium grown on PDA plates.

	Number of differentially expressed genes (96 hpi / mycelium)		
	FOP1 EMR	**F81**	**R2**
Upregulated	3492	3871	3971
Downregulated	4455	4259	4500
Total	7947	8130	8471

FOP genes that were considered differentially expressed (determined as above) were arranged in descending order based on the highest expression *in planta* (upregulated) and *in vitro* (downregulated). The top 30 genes from each list for each isolate (FOP1 EMR, F81 and R2) were examined, along with additional annotations from the genome analyses (Chapter 3), to determine potentially important genes in FOP during infection.

For FOP1 EMR (race 1), the 30 highest expressed upregulated genes included 10 which were predicted to be secreted (as they had a signal peptide), with two of these being within 2 kb of a mimp (Table 4.4), which was indicative of an effector gene (Chapter 3). Of these 30 genes, 13 were identified as putative effectors (using EffectorP), including the top two highest expressed genes (g16048, g17037). Four genes were identified as CAZymes (3 secreted; g8487, g8835, g18598) with predicted functions such as glucanase activity, chitinase activity and reductase-like activity. The fifth highest expressed gene (g18161) in FOP1 EMR, on contig 24, was predicted to both be secreted and an effector

protein with two orthologs in this isolate with only one ortholog present in the other two FOP isolates, and none identified in FOC or FOL. One gene (g20448) was recognised by interproscan and swissprot as encoding pisatin demethylase cytochrome P450 enzymes that are commonly associated with *F. solani*. The 30 most highly downregulated genes in FOP1 EMR included only four that were predicted as secreted, with none of these being within 2 kb of a mimp or being identified as a CAZyme. Functions associated with these genes through interproscan included immunoglobulin, stress tolerance transmembrane domain, chromatin organisation modifier and sugar transporters (Table A 4.2).

For F81 (race 2), the 30 most highly expressed upregulated genes included seven predicted as secreted, although none were predicted within 2 kb of a mimp or recognised as CAZymes (Table 4.5). EffectorP predicted 11 genes as putative effectors with many identified as ribosomal proteins. The 20th highest expressed gene was identified as *SIX1*, located on contig 66 (LS, Chapter 3), and was identified as orthologous between F81, R2 and FOL. Other functions predicted within these genes included serine proteases (trypsin) and ubiquitin domains. Investigating the most downregulated genes for F81, seven were found to be secreted, and two were identified as transmembrane domains. Additional genes contained interproscan annotations suggesting they are membrane bound proteins, as well as other genes annotated as mRNA binding proteins and cysteine rich secreted proteins (Table A 4.3).

The results for the most highly expressed genes in R2 (race 5) were similar to F81, in that seven were predicted as secreted proteins, and none were within 2 kb of a mimp. One gene (g6883) was predicted as a CAZyme (CAZY:PL3) with associated pectate lyase function. Again, *SIX1* was among the most highly expressed genes in R2 (7th place), and was the same copy that was highly expressed in F81. There were nine genes predicted by EffectorP as being putative effectors, one of which (g5964) was on contig 4 (same contig as the CAZy gene), and which was orthologous across all FOP genomes as well as FOC and FOL. This was identified by PHIbase as an avirulence determinant (Table 4.6). Downregulated genes in R2 included two that were predicted as secreted (g4113 and g4213), two predicted as transmembrane domains (g201 and g14224) and three transcription factors. Five genes showed no orthology with F81 including one (g4113) which was predicted to be secreted and also identified as a putative effector by EffectorP (Table A 4.4).

Table 4.4 The 30 most highly expressed upregulated genes following RNAseq analysis of pea roots infected with *Fusarium oxysporum* f. sp. *pisi* (FOP) isolate FOP1 EMR (race 1) at 96 hpi compared to *in vitro* grown mycelium. Expression levels are ranked in descending order using the average FPKM (fragments per kilobase million) values from three replicates for the 96 hpi data. Genome annotations are details from interproscan which suggest possible associated protein signatures, along with any predicted carbohydrate active enzymes (CAZyme) identified from the CAZy database. Predicted secreted and effector-like proteins (EffP), and numbers of proteins identified as orthologous are included. Orthogroup counts are between the isolates F81(race 2)**:**R2(race 5)**:**FOP1EMR(race 1)**:**FOC**:**FOL.

Gene name	Contig	Secreted	EffP	Orthogroup contents	Gene annotations	FPKM (96 h)	FPKM (myc.)
g16048	16	Yes	Yes	0:0:1:1:0	2 kb of a mimp	37033	3507
g17037	19	Yes	Yes			16932	1774
g7850	5			1:1:1:1:1	Ubiquitin related domain	8254	2423
g10391	7			9:9:11:9:20	Dehydrogenase related	7871	1278
g18161	24	Yes	Yes	1:1:2:0:0		6332	105
g20446	38			3:2:4:1:1	Polyketide cyclase	5347	1
g5956	4	Yes	Yes	1:1:1:1:1		4825	39
g9964	6			4:4:4:1:2	NADP binding domain	4613	12
g2788	2		Yes	1:2:1:1:1	Nitroreductase family	4553	26
g1109	1			1:1:1:2:1	Ser-Thr-rich membrane family	4064	1586
g11581	8		Yes	1:1:1:1:1	Polyketide cyclase	3689	2
g17396	20			4:5:4:2:3	Dehydrogenase related	3623	1
g1217	1	Yes		1:1:1:1:1	Ser-Thr-rich membrane family	3565	1480
g19310	29			4:5:4:2:3	Dehydrogenase related	3440	1
g10965	7		Yes	1:1:1:1:1	Cupin domain	3324	0
g8487	5	Yes		1:1:1:1:0	Splicing factor; CAZY:GH131	3310	1215
g8835	6	Yes		5:4:6:4:5	Chitin binding chitinase; CAZY:GH18	3152	85
g4578	3			1:1:1:1:1		3052	815
g7902	5		Yes	2:1:1:1:1	Zinc superoxide dismutase	2972	903
g18598	26	Yes		1:1:1:1:1	Glycosyl hydrolases family; CAZY:GH18	2838	628
g10961	7		Yes	9:6:5:6:5	NAD(P)-binding domain	2821	0
g17011	19	Yes		1:1:1:1:1	2 kb of a mimp	2631	679
g12835	9	Yes		1:1:1:1:2		2533	113
g508	1			3:3:3:3:6	ATPases phosphorylation site	2291	808
g11586	8		Yes	1:1:1:1:1	Stress responsive alpha-beta barrel	2274	3
g10966	7		Yes	1:1:1:1:1		2232	13
g10486	7		Yes	1:1:1:1:1	Formaldehyde-activating enzyme/centromere protein	2086	263
g20448	38			32:27:30:26:22	Cytochrome P450 cysteine ligand (Pisatin demethylase)	2082	1
g533	1		Yes	1:1:1:1:1	NADPH-dependent FMN reductase; CAZY:AA6	2011	91
g20447	38			1:1:1:0:0	NmrA-like family	1974	0

Table 4.5 The 30 most highly expressed upregulated genes following RNAseq analysis of pea roots infected with *Fusarium oxysporum* f. sp. *pisi* (FOP) isolate F81 (race 2) at 96 hpi compared to *in vitro* grown mycelium. Expression levels are ranked in descending order using the average FPKM (fragments per kilobase million) values from three replicates for the 96 hpi data. Genome annotations are details from interproscan which suggest possible associated protein signatures, along with any predicted carbohydrate active enzymes (CAZyme) identified from the CAZy database. Predicted secreted and effector-like proteins (EffP), and numbers of proteins identified as orthologous are included. F81(race 2)**:**R2(race 5)**:**FOP1EMR(race 1)**:**FOC**:**FOL.

Gene name	Contig	Secreted	EffP	Orthogroup contents	Gene annotations	FPKM (96 h)	FPKM (myc.)
g16856	21			2:2:1:2:3		53229	23647
g20182	58		Yes	1:1:1:1:0		21277	9448
g3390	3				Ribosomal protein L39e	14927	7082
g12483	12			1:1:1:1:1		14837	6586
g6196	5	Yes		1:1:1:1:1	Serine proteases	14081	123
g6841	6	Yes		1:1:1:1:1	Splicing factor	12863	3290
g3092	3		Yes	1:1:1:1:1	Ribosomal protein S14	10796	5319
g9823	9	Yes	Yes	1:1:1:1:0	CFEM domain	9337	103
g14175	14		Yes	2:2:2:0:0	S25 ribosomal protein	9294	4292
g16557	20		Yes	1:1:1:1:1		8246	3941
g7098	6			1:1:1:1:1	Ribosomal protein L12 family	8071	3224
g12543	12			1:1:1:1:1	Ribosomal protein P2	7892	3573
g12294	12			1:1:1:1:1	Ribosomal protein L29e	7507	3030
g1898	2			1:2:1:2:1	Ubiquitin domain	7463	2732
g1721	2			1:1:1:1:1	phospho-glucose isomerase	6781	27
g244	1		Yes	1:1:1:1:1	Ribosomal protein S21e	6648	3242
g3560	3		Yes	1:1:1:1:1	ribosomal protein L37Ae	6068	2505
g8124	7			1:1:1:1:1	Peptidase inhibitor I9	6065	2697
g6575	6			2:3:0:0:1	Cytochrome c oxidase	5882	2810
g20369	66	Yes		1:1:1:1:1	Secreted In Xylem 1 (*SIX1*^{b})	5800	17
g2007	2			1:1:1:1:1	50S ribosomal protein L37e	5750	2545
g14174	14			1:1:2:0:0	Mitochondrial carrier domain	5471	1942
g19265	40	Yes	Yes	1:1:1:1:1		5378	34
g12425	12		Yes	1:1:1:1:1	Ribosomal protein S10 domain	5342	2244
g6895	6			1:1:1:1:1	Ribosomal protein L44	4960	2119
g1354	1		Yes	1:1:1:1:1	Ribosomal protein L35Ae	4742	2143
g741	1	Yes		1:1:1:1:4	Ser-Thr-rich membrane family	4710	1379
g15191	16			1:1:0:1:1	Ubiquitin domain	4615	2245
g18997	36	Yes	Yes	9:9:11:9:20		4584	3
g15289	16			2:2:1:2:3	Dehydrogenase	4518	465

Table 4.6 The 30 most highly expressed upregulated genes following RNAseq analysis of pea roots infected with *Fusarium oxysporum* f. sp. *pisi* (FOP) isolate R2 (race 5) at 96 hpi compared to *in vitro* grown mycelium. Expression levels are ranked in descending order using the average FPKM (fragments per kilobase million) values from three replicates for the 96 hpi data. Genome annotations are details from interproscan which suggest possible associated protein signatures, along with any predicted carbohydrate active enzymes (CAZyme) identified from the CAZy database. Predicted secreted and effector-like proteins (EffP), and numbers of proteins identified as orthologous are included. Orthogroup counts are between the isolates F81(race 2)**:**R2(race 5)**:**FOP1EMR(race 1)**:**FOC**:**FOL.

Gene name	Contig	Secreted	EffP	Orthogroup contents	Gene annotations	FPKM (96 h)	FPKM (myc.)
g7324	5			9:9:11:9:20	Dehydrogenase family	15229	1880
g8973	6			1:1:1:1:1	Ubiquitin domain	13033	2428
g15550	13		Yes			12029	4085
g670	1			58:53:44:47:53	Aldehyde dehydrogenase	10122	1281
g18205	21	Yes	Yes	1:1:0:0:0		9941	281
g391	1	Yes	Yes	1:1:1:1:1		9523	3546
g17379	18	Yes		2:3:0:0:1	Secreted In Xylem 1 (*SIX1*^{b})	9349	3
g2848	2		Yes	1:1:1:1:1	Ribosomal protein S14	9295	4137
g11073	8			1:1:1:1:1		7986	1259
g14263	11			1:1:1:1:2	ATP synthase subunit c	7543	1244
g14235	11			1:1:1:1:1	Mitochondrial carrier protein	7042	1628
g6121	4	Yes		1:1:1:1:1	Splicing factor	7004	991
g6388	4			1:1:1:1:1	Cytochrome c oxidase	6602	2717
g7743	5			1:1:1:1:1	Ribosomal protein P2	6541	3066
g12283	9			1:1:1:1:1	Cytochrome c oxidase	6305	2776
g5964	4	Yes	Yes	1:1:1:1:1	Cerato-platanin	6136	438
g5977	4			1:1:1:1:2	Membrane protein	5898	78
g5862	4			1:1:1:1:1	ribosomal protein L12	5547	2588
g8468	6			3:3:3:3:5	translation elongation factor EF-1, subunit alpha	5527	2504
g9106	6			3:3:3:3:3	Splicing factor	5226	368
g7494	5			1:1:1:1:1	Ribosomal protein L29e	5031	2257
g6883	4	Yes	Yes	1:1:1:1:1	Pectin lyase fold/virulence factor	4877	1
g17587	18			2:1:3:0:0	Galactose-binding domain/dipeptidyl-peptidase	4765	32
g3317	2		Yes	1:1:1:1:1	Ribosomal protein L37ae	4742	1992
g8863	6			1:1:1:1:1	Ribosomal protein L37e	4429	2086
g7298	5		Yes	1:1:2:1:1	Class IA and IB cytochrome C signature	4423	548
g17468	18					4381	488
g15580	13		Yes			4295	207
g16698	15	Yes		1:1:1:1:1	Splicing factor	4263	301
g14350	11			1:1:1:1:15	D-isomer specific 2-hydroxyacid dehydrogenase	4233	577

4.3.3.3 Assessing *SIX* gene differential expression in FOP

Of the eight *SIX* genes in FOP isolate FOP1 EMR (race 1, *SIX1*, *SIX6*, *SIX7*, *SIX9*, *SIX10*, *SIX11*, *SIX12* and *SIX14*), only *SIX1*, *SIX7*, *SIX10* and *SIX12* were classified as being upregulated (using DESeq2) and were ranked at 1118th, 556th, 1374th and 561st place respectively out of the total of 3492 upregulated genes (Table 4.7). Average gene expression values for *SIX11* might suggest upregulation at 96 hpi; however, there was large variation between replicates and this therefore was not significantly different (data not shown). One of the copies of *SIX6* and the *SIX10* gene both resided on the same contig (44, LS Chapter 3), which was also true for *SIX7* and *SIX12* (contig 59, LS Chapter 3). Six of the eight *SIX* genes in FOP1 EMR were identified as putative effector proteins using EffectorP, and three were located within 2 kb of a mimp (Chapter 3). One gene (*SIX7*) was recognised as a transmembrane domain, and therefore was not predicted to be secreted. Both *SIX9* and *SIX14* were not found during the gene prediction stage in the genome analysis (Chapter 3) and therefore could not be included in the RNAseq analysis.

All five of the *SIX* genes present in FOP isolate F81 (race 2, *SIX1*, *SIX6*, *SIX9*, *SIX13*, *SIX14*) were upregulated during infection compared to the control mycelium samples. When the most highly expressed upregulated genes were ranked, *SIX1* was at 20th place, *SIX14* was at 385th place, *SIX6* was 533rd, *SIX13* was 1183rd and *SIX9* was 2109th place out of the total of 3878 upregulated genes. The extra copy of *SIX1* also found in F81 was also expressed at a low level (2945th), and was located within 2 kb of a mimp (Table 4.7). *SIX6* and *SIX13* were also found to be located within 2 kb of a mimp. *SIX6*, *SIX9* and *SIX14* were recognised as putative effectors using EffectorP, and all but one *SIX* gene (*SIX13*) was predicted as being secreted (Chapter 3).

For the three *SIX* genes present in FOP isolate R2 (race 5, *SIX1*, *SIX6*, *SIX13*) the most highly expressed was a copy of *SIX1* ranked at 7th place out of the total of 3871 upregulated genes. Another copy of *SIX1* (not present in F81) was also upregulated during infection and was ranked 1929th. A third copy (matching to g18035 in F81) was not expressed during infection. *SIX6* and *SIX13* present in R2 were also upregulated during infection and were ranked at 722nd and 2171st respectively out of the 3971 genes upregulated. All *SIX* genes were predicted to be secreted, with all but the highest expressed copy of *SIX1* being located within 2 kb of a mimp (Table 4.7). *SIX6* was recognised as a putative effector by EffectorP, and was the only gene located on a separate

contig (21, LS Chapter 3) compared to all the others being located on contig 18 (LS, Chapter 3).

Table 4.7 Location, predicted secretion, gene annotation and relative expression (using average FPKM values) for *Secreted In Xylem* (*SIX* genes) identified in the genomes of *Fusarium oxysporum* f. sp. *pisi* (FOP) isolates FOP1 EMR (race 1), F81 (race 2) and R2 (race 5), based on RNAseq analysis of infected pea roots (96 hpi) and for *in vitro* grown mycelium. Genes were classified as upregulated (differentially expressed) or not based on DESeq2 analysis.

Isolate	Gene name	*SIX* gene	Contig	Secreted	Gene annotations	FPKM (96 h)	FPKM (myc.)	DEG cond.
F1 EMR	g20529	*SIX1*[1]	39		EffP	24	1	Upreg
F1 EMR	g17863	*SIX6*[2]	22	Yes	EffP	10	2	
F1 EMR	g20948	*SIX6*[1]	44	Yes	EffP	0	0	
F1 EMR	g21760	*SIX7*	59		2 kb mimp, TMD	100	1	Upreg
F1 EMR	g20977	*SIX10*	44	Yes	2 kb mimp, EffP	12	0	Upreg
F1 EMR	g21938	*SIX11*	66	Yes	EffP	1524	525	
F1 EMR	g21759	*SIX12*	59	Yes	2 kb mimp, EffP	99	1	Upreg
F81	g20369	*SIX1*[2]	66	Yes		5800	17	Upreg
F81	g18035	*SIX1*[3]	28	Yes	2 kb mimp	5	0	Upreg
F81	g20417	*SIX6*[2]	69	Yes	2 kb mimp, EffP	250	0	Upreg
F81	g8636	*SIX9*	8	Yes	EffP	20	0	Upreg
F81	g17065	*SIX13*	22		2 kb mimp	76	2	Upreg
F81	g18445	*SIX14*	31	Yes	EffP	370	62	Upreg
R2	g17379	*SIX1*[2]	18	Yes		9349	3	Upreg
R2	g17417	*SIX1*[3]	18	Yes	2 kb mimp	32	0	Upreg
R2	g17623	*SIX1*[4]	18	Yes	2 kb mimp	0	0	
R2	g18204	*SIX6*[2]	21	Yes	2 kb mimp, EffP	169	0	Upreg
R2	g17466	*SIX13*	18	Yes	2 kb mimp	23	1	Upreg

TMD = Transmembrane domain
EffP = Predicted effector protein using EffectorP
2 kb mimp = gene was located within 2 kb of a mimp
F1 EMR = FOP1 EMR
Upreg = Upregulated assigned by DESeq

4.3.3.4 Prediction of additional putative effector genes in FOP

Although *SIX* genes were present and the majority were expressed in the different FOP isolates, they may not be the only genes associated with pathogenicity. Therefore, predicted secreted genes located within 2 kb of a mimp were analysed for each isolate (FOP1 EMR, F81 and R2), following the approach taken by Schmidt *et al.* (2013), (see Section 4.1 Introduction and Chapter 3) as a means of identifying novel effectors, especially those which were highly expressed. Between isolates and therefore races (1, 2 and 5), there were examples of shared genes, based on orthology analysis and functional annotation (Chapter 3). In FOP1 EMR (race 1), 24 secreted genes were within 2 kb of a mimp (Table 4.8), with 17 genes identified in F81 (race 2; Table 4.9) and 21 genes in R2 (race 5; Table 4.10). Of these secreted genes within 2 kb of a mimp, there was one gene

(g17046) identified in FOP1 EMR, three genes (g18880, g18891, g20417) in F81 and three genes g17990, g18001, g18184) in R2 that shared orthogroups with genes in the other FOP races. These genes were identified as putative effectors unique to FOP, potentially important for infection as they were all highly expressed *in planta*, and were also identified in the orthology analyses in Chapter 3.

There were nine genes (FOP1 EMR and F81) and 12 genes (R2) predicted as secreted and within 2 kb of a mimp which were in shared orthogroups with other FOP isolates, FOC and FOL. Functional annotations for these genes included concanavalin A-like lectins, alginate lyases and many CAZymes belonging to carbohydrate binding module families, auxiliary activity families and glycoside hydrolase families. These shared genes were differentially expressed *in planta*, with some genes coding for LysM domains, concanavalin A-like lectins and alginate lyases being highly expressed in all FOP isolates (Table 4.8, Table 4.9 and Table 4.10). There were further genes shared between F81 (race 2), R2 (race 5) and FOL including *SIX1*, *SIX6* and *SIX13* which were all upregulated during infection.

In all three FOP isolates there were examples of unique secreted genes within 2 kb of a mimp, which were highly expressed during infection, and are likely to represent putative race specific effectors. Three genes (g21410, g21926, g21927) were identified as unique to FOP1 EMR (race 1), all of which were highly upregulated during infection and also identified as putative effectors by EffectorP (Table 4.8). There were no unique genes specific to F81 (race 2) and R2 (race 5), however there were two highly expressed genes in F81 (g17126, g19532) and R2 (g17520, g17419) shared exclusively between these isolates, which were also identified as putative effectors by EffectorP (Table 4.9 and Table 4.10).

Table 4.8 Details of 24 genes predicted as secreted and located within 2 kb of a mimp, present in *Fusarium oxysporum* f. sp. *pisi* (FOP) isolate FOP1 EMR determined by genome analysis, and their associated normalised expression values (FPKM) from infected pea roots (96 hpi) and mycelium (myc.). Genome annotations are details from interproscan which suggest possible associated protein signatures, along with any predicted carbohydrate active enzymes (CAZyme) identified by the CAZY database. Predicted effector-like structures (EffP) and numbers of proteins identified as orthologous are included. Orthogroup counts are between the isolates F81(race 2)**:**R2(race 5)**:**FOP1EMR(race 1)**:**FOC**:**FOL. DE (differential expression) status is shown if it is upregulated.

Gene	Contig	EffP	Orthogroup contents	Gene annotations	FPKM (96 h)	FPKM (myc.)	DE
g16048	16	Yes	0:0:1:1:0		37033	3507	up
g16054	16	Yes	0:0:1:1:0		4364	148	
g21893	64	Yes	2:2:2:0:1		3415	11	
g17011	19		1:1:1:1:1		2631	679	up
g20246	36		2:3:3:2:2	Concanavalin A-like lectin	1892	63	up
g20255	36		2:3:3:2:2	Concanavalin A-like lectin	1892	63	up
g17056	19	Yes	0:0:1:1:0		1692	7	up
g17034	19	Yes	1:1:2:1:0		888	9	up
g21206	47		4:5:4:3:3	Alginate lyase domain	618	4	
g19441	30	Yes	1:1:3:1:1	LysM domain, CAZY:CBM50	603	1	up
g21410	51	Yes	0:0:4:0:0		440	0	up
g21926	65	Yes	0:0:4:0:0		440	0	up
g21927	65	Yes	0:0:4:0:0		440	0	up
g20952	44				328	34	up
g17046	19	Yes	2:3:2:0:0		175	1	up
g20992	44		0:0:3:2:0	Peptidase, metallopeptidase	148	4	
g20477	39		13:10:11:8:12	Glucose-methanol-choline oxidoreductase, CAZY:AA3	122	0	up
g21759	59	Yes	0:0:1:1:0		99	1	up
g20253	36		6:6:5:3:7	Plant heme peroxidase, CAZY:AA2	73	0	up
g19587	31	Yes			26	0	up
g20977	44	Yes	0:0:1:1:1		12	0	up
g17883	22		5:7:20:3:11	Fungal specific transcription factor	5	1	up
g21468	51		0:0:3:2:0	Peptidase, metallopeptidase	1	0	
g20533	39		6:6:5:3:7	Plant heme peroxidase, CAZY:AA2	1	0	up

Table 4.9 Details of 17 genes predicted as secreted and located within 2 kb of a mimp, present in *Fusarium oxysporum* f. sp. *pisi* (FOP) isolate F81determined by genome analysis, and their associated normalised expression values (FPKM) from infected pea roots (96 hpi) and mycelium (myc.). Genome annotations are details from interproscan which suggest possible associated protein signatures, along with any predicted carbohydrate active enzymes (CAZyme) identified by the CAZY database. Predicted effector-like structures (EffP) and numbers of proteins identified as orthologous are included. Orthogroup counts are between the isolates F81:R2:FOP1EMR:FOC:FOL.

Gene	Contig	EffP	Orthogroup contents	Gene annotations	FPKM (96 h)	FPKM (myc.)	DE
g19728	46	Yes	1:1:3:1:1	LysM domain; CAZY:CBM50	1105	6	up
g17126	22	Yes	2:1:0:0:0		952	197	up
g18880	35	Yes	3:3:6:0:0		704	1	up
g18891	35	Yes	3:3:6:0:0		704	1	up
g19747	46	Yes	1:1:1:0:0		452	1	up
g19292	40		2:3:3:2:2	Concanavalin A-like lectin	410	162	up
g17049	22		4:5:4:3:3	Alginate lyase domain	308	2	up
g20417	69	Yes	1:1:2:0:1	*SIX6*[2]	250	0	up
g19532	43	Yes	1:1:0:0:0		203	0	up
g18041	28		13:10:11:8:12	Glucose-methanol-choline oxidoreductase; CAZY:AA3	118	0	up
g16409	19		3:2:4:4:2	LysM domain, CAZY:CBM50	73	43	
g17064	22	Yes	3:1:0:0:2		71	2	up
g18035	28		2:3:0:0:1	*SIX1*[3]	5	0	up
g19527	43		157:130:604:114:29	Ribonuclease H-like domain	5	0	up
g17056	22		2:3:1:1:2	Short-chain dehydrogenase	4	1	up
g17966	27		13:12:7:13:16	Fibronectin type III-like domain; CAZY:GH3	3	0	up
g19775	47		3:1:2:2:2		0	0	

Table 4.10 Details of 21 genes predicted as secreted and located within 2 kb of a mimp, present in *Fusarium oxysporum* f. sp. *pisi* (FOP) isolate R2 determined by genome analysis, and their associated normalised expression values (FPKM) from infected pea roots (96 hpi) and mycelium (myc.). Genome annotations are details from interproscan which suggest possible associated protein signatures, along with any predicted carbohydrate active enzymes (CAZyme) identified by the CAZY database. Predicted effector-like structures (EffP) and numbers of proteins identified as orthologous are included. Orthogroup counts are between the isolates F81(race 2)**:**R2(race 5)**:**FOP1EMR(race 1)**:**FOC**:**FOL.

Gene	Contig	EffP	Orthogroup contents	Gene annotations	FPKM (96 h)	FPKM (myc.)	DE
g17520	18	Yes	2:1:0:0:0		2596	448	up
g18165	21	Yes	1:1:3:1:1	LysM domain; CAZY:CBM50	889	14	up
g18236	21		4:5:4:3:3	Alginate lyase domain	605	1	up
g17451	18		4:5:4:3:3	Alginate lyase domain	546	0	up
g17990	20	Yes	3:3:6:0:0		469	0	up
g18001	20	Yes	3:3:6:0:0		469	0	up
g17617	18		13:10:11:8:12	Glucose-methanol-choline oxidoreductase; CAZY:AA3	365	0	up
g18184	21	Yes	1:1:1:0:0		316	0	up
g17920	20		2:3:3:2:2	Concanavalin A-like lectin	205	139	
g18204	21	Yes	1:1:2:0:1	*SIX6*[2]	169	0	up
g17053	16	Yes	1:2:3:2:2		165	3	up
g16998	16		9:8:7:9:11	PCMH-type FAD-binding domain; CAZY:AA7	68	1	up
g17419	18	Yes	1:2:0:0:0		62	9	up
g17417	18		2:3:0:0:1	*SIX1*[3]	32	0	up
g17466	18		3:1:0:0:2	*SIX13*	23	1	up
g17459	18		2:3:1:1:2	Short-chain dehydrogenase	13	0	up
g18212	21		6:6:5:3:7	Haem peroxidase; CAZY:AA2	12	0	up
g18104	21		3:3:1:1:1	Pectin lyase fold; CAZY:GH28	6	0	up
g16251	14		157:130:604:114:29	Ribonuclease H-like domain	1	0	up
g17623	18		2:3:0:0:1	*SIX1*[4]	0	0	
g16343	15		2:2:1:2:2	Glycoside hydrolase; CAZY:GH5	0	0	

4.4 Discussion

In this chapter, *SIX* genes were found in FOP isolates representing three different races but were mostly absent in the *F. oxysporum* root rot isolates obtained from UK fields. An extensive screen of FOP isolates using PCR confirmed that there was race dependent distributions of *SIX* genes as hypothesised in Chapter 3 from genome sequencing. Race 1 was confirmed to contain *SIX1*, *SIX6*, *SIX7*, *SIX9*, *SIX10*, *SIX11*, *SIX12* and *SIX14*, race 2 was confirmed to contain *SIX1*, *SIX6*, *SIX9*, *SIX13*, *SIX14* while race 5 was confirmed to contain *SIX1*, *SIX6*, *SIX13*. To our knowledge this is the first screen of large numbers of FOP isolates for *SIX* gene presence, which has confirmed that different races can be distinguished in this way. This study also revealed, using qPCR analysis, that most of the *SIX* genes are expressed *in planta* and that expression levels increased over time during the early stages of infection. RNAseq analysis has revealed that additional putative effectors (identified in this chapter and in Chapter 3) were highly expressed *in planta* and could therefore be important for pathogenicity of FOP in pea. The implication of a race specific complement of *SIX* genes indicates targets to molecularly distinguish between races, avoiding the need for time-consuming plant assays.

Many of the previously race typed FOP isolates (mainly from the USA, Chapter 2) retained their original race designation according to their complement of *SIX* genes as determined by PCR. However, a small selection of FOP isolates which were previously identified as race 2 lacked *SIX9* and *SIX14* (found in all other race 2 isolates) and matched the *SIX* gene complement of the historic race 5 isolate FOP5 (*SIX1*, *SIX6*, *SIX13*). These were therefore reclassified as race 5 isolates, which was also supported by their grouping into the same phylogenetic *TEF* clade as FOP5 (Section 2.3.2). There was a distinct difference of five *SIX* genes between FOP races 1 (*SIX1*, *SIX6*, *SIX7*, *SIX9*, *SIX10*, *SIX11*, *SIX12* and *SIX14*) and 2 (*SIX1*, *SIX6*, *SIX9*, *SIX13*, *SIX14*), which also separated into separate phylogenetic *TEF* clades (Section 2.3.2 and 4.3.1). Differences in *SIX* gene profiles between races has been reported in other f. spp. of *F. oxysporum*, and it has been suggested as a method to molecularly distinguish races. In FOL, the presence of *SIX4* only in race 1 isolates has been used to distinguish it from races 2 and 3 where it is absent (Lievens *et al.*, 2009). In addition, differences in *SIX8* gene sequence can be used to distinguish race 4 from races 1 and 2 in *F. oxysporum* f. sp. *cubense* (Fraser-Smith *et al.*, 2014). In the current study, the complement of *SIX* genes in putative FOP races has been

determined, which has identified gene targets that could be used as means of distinguishing FOP races.

A *SIX* gene race structure in FOP was briefly alluded to by Taylor *et al.* (2016) after screening of isolates FOP1 (race 1), FOP2 (race 2) and FOP5 (race 5) with mostly FOC specific and FOL specific *SIX* gene PCR primers. They found that FOP1 contained five *SIX* genes (*SIX7*, *SIX10*, *SIX11*, *SIX12* and *SIX14*), FOP2 contained two *SIX* genes (*SIX13* and *SIX14*) and FOP5 contained one *SIX* gene (*SIX13*) and suggested a distinct variation between races in terms of *SIX* gene complement (Taylor *et al.*, 2016). In contrast to the current study where FOP specific *SIX6* primers were designed and used, Taylor *et al.* (2016) did not identify *SIX6* in any of the three FOP races. Sequence variation in *SIX6* between different *F. oxysporum* f. spp. has been widely reported, with one study finding only 60% identity between *SIX6* in *F. oxysporum* f. sp. *cubense* and FOL for example (Gawehns *et al.*, 2014), which may explain why the same primers are not always successful for different f. spp. Genome analysis of a FOP race 5 isolate identified *SIX1*, *SIX9*, *SIX13* and *SIX14* (Williams *et al.*, 2016) which differs from the results in this study where only *SIX1*, *SIX6* and *SIX13* were identified. This discrepancy in results could be due to incorrect race typing of isolates, or that there is variation in *SIX* gene complements within the same races.

F. oxysporum isolates collected from diseased pea plants in UK fields showing symptoms of root rot or wilt (Chapter 2) were also screened for the presence of *SIX* genes with the FOP specific PCR primers. As they contained few or no *SIX* genes they were designated root rot isolates rather than wilt-causing FOP isolates. A small number of *F. oxysporum* isolates from the USA, Czech Republic and Algeria previously identified as FOP based on pathogenicity tests were also reclassified as root rot isolates, as they had few or no *SIX* genes. Therefore, it is probable that they were previously incorrectly race typed, as it is not clear what effect some root rot isolates would have on pea differential cultivars during a race typing experiment. These *SIX* genes may have been acquired in root rot *F. oxysporum* isolates by horizontal transfer, however movement of only part of the LS regions could result in the gain of only one or two genes. It is possible that *SIX* genes in root rot isolates may not contribute to pathogenicity and that more may be required for full virulence. In FOL, *SIX3* and *SIX5* are both required to overcome I-2-mediated disease

resistance in tomato, and have been shown to interact at plasmodesmata sites (Cao *et al.*, 2018).

SIX genes were upregulated during the early stages of pea infection in all FOP isolates tested using qPCR, however, very different levels of expression were observed between genes. The relative expression levels observed in F81 (race 2) and R2 (race 5) for *SIX1*^{2} and *SIX13* differed by a factor of approximately 1000, with expression levels in *SIX1* exceeding expression levels of all other *SIX* genes. This suggests that *SIX1*^{2} may have a greater role in the pathogenicity of FOP than other *SIX* genes over the time measured. *SIX1* has previously been shown to be essential for virulence in FOL (Rep *et al.*, 2005). Due to the multiple copies of *SIX1* identified, primers were designed to be specific to the highest expressed copy found in both F81 (race 2) and R2 (race 5) (*SIX1*^{2}), which was not present in FOP1 EMR (race 1) and therefore no qPCR expression data could be obtained for this isolate. *SIX1* sequences have been shown to vary between races of *F. oxysporum* f. sp. *cubense*, with the sequence of *SIX1* in sub-tropical race 4 differing from the sequence found in race 1, 2 and tropical race 4 (Meldrum *et al.*, 2012). The variation in sequence could have allowed sub-tropical race 4 to adapt to the cooler temperatures required to proliferate in sub-tropical regions (Ploetz, 2015).

SIX6 was also expressed in all FOP isolates during infection of pea roots, which may suggest that it is also involved in pathogenicity. *SIX6* has been shown to be necessary for full virulence in FOL (Gawehns *et al.*, 2014) which may also be true for FOP based on its expression *in planta*. There was variation in the amount of expression of *SIX6* between isolates which could be due to experimental conditions such as differences in growth rates between isolates or the aggressiveness of isolates in an artificial inoculation system. Only one copy of *SIX6* (*SIX6*^{2}) was used for qPCR analysis due to the similarities in sequences between *SIX6*^{1} and *SIX6*^{2} in FOP1 EMR. The additional copy of *SIX6* in FOP1 EMR could have resulted from duplication and mutation of the *SIX6*^{1} gene in order for FOP race 1 to maintain pathogenicity in pea. Although *SIX9* and *SIX14* were present in the genome of FOP1 EMR, and were detected via PCR screening, they were not highly expressed *in planta*, yielding expression values lower than the limit of detection, suggesting they could be pseudogenes and therefore not important in infection.

SIX gene expression in FOP was also examined using RNAseq data which allowed expression levels to be quantified relative to other highly expressed genes *in planta*. In

FOP1 EMR (race 1) only four *SIX* genes (*SIX1*[1], *SIX7*, *SIX10* and *SIX12*) out of the total identified were classified as upregulated by DESeq2 using the RNAseq data, which could be due to variation in expression levels between biological replicates (*SIX11* and *SIX6*[2]) or no expression detected (*SIX6*[1]). *SIX7* was predicted to have a transmembrane domain and therefore is not likely to be a secreted cytoplasmic effector, unlike for other f. spp. where *SIX7* has been identified in xylem sap of infected tomatoes (Lievens *et al.*, 2009). In both FOP isolates F81 (race 2) and R2 (race 5), *SIX1*[2] was highly upregulated during infection, although this was the only copy in these isolates which was not within 2 kb of a mimp. High expression suggests that this copy is important for pathogenicity in race 2 and 5 isolates, as it was also one of the 30 highest expressed upregulated genes. A version of *SIX1* was upregulated during infection in all isolates suggesting that it could also be essential for full virulence in FOP.

The *SIX* gene complement and expression in FOP races could help to explain the evolution of races. Based on *TEF* phylogeny (Section 2.3.2) race 1 and race 2 are closely related, therefore the potential loss of *SIX7*, *SIX10*, *SIX11* and *SIX12* from race 1 and the gain of *SIX13* could have resulted in the emergence of race 2. Loss of any of these genes could have enabled race 2 to evade specific host defence responses. These *SIX* genes may not be important for pathogenicity and may be redundant even in race 1, which could have also contributed to their loss during the evolution of a new race. Putative race 5 isolates were hypothesised to be from a different lineage based on *TEF* phylogeny (Chapter 2), which is supported by its complement of *SIX* genes. The only shared and highly expressed *SIX* genes in all three races were *SIX1* and *SIX6*, therefore it is possible that these are essential for pathogenicity of FOP in pea.

The 30 highest expressed upregulated genes in all three FOP isolates included carbohydrate active enzymes (CAZymes) of the glycoside hydrolase, lignin-degrading and auxiliary activity families, which are all important during fungal infection as they degrade plant cell walls (Ospina-Giraldo *et al.*, 2010). In addition, a gene (g20448, FOP1 EMR) was also identified with cytochrome P450 activity, noted as similar to Pisatin demethylase in *F. solani*, which is reported to be involved in pathogenicity of *F. solani* and more recently of *F. oxysporum* in pea (Coleman *et al.*, 2011). Other highly expressed genes included serine proteases (thought to be involved in host colonisation, Olivieri *et al.*, 2002) and CFEM domains (highly enriched in pathogenic fungi and involved in

pathogenicity of rice blast fungus *Magnaporthe grisea*, Zhang *et al.*, 2015) and therefore could also be associated with pathogenicity. Genes such as ribosomal proteins and cytochrome oxidase genes (involved in protein translation and secondary metabolism, Schmidt *et al.*, 2013), pectin lyases (which degrades pectin found in the middle lamella between plant cell walls, Yadav *et al.*, 2017) and cerato-platanin (induces host defence responses in plants, Gomes *et al.*, 2015) were also featured in the 30 highest expressed genes *in planta*. These results agree with findings from similar studies where expression of genes related to pathogenicity, such as cytochrome p450s, glycoside hydrolases, peptidases and fungal transcription factors were reported for FOP, *F. oxysporum* f. sp. *ciceris* and f. sp. *medicaginis* (Williams *et al.*, 2016; Thatcher *et al.*, 2016).

Expression levels of putative effectors identified as having a signal for secretion and being located within 2 kb of a mimp were analysed for genes orthologous between FOP, FOC and FOL. Highly expressed orthologous genes included concanavalin A-like lectins (facilitate penetration of parasitic fungi into the host, Khan & Khan, 2011), alginate lyase domains (a type of CAZyme, so involved in degradation of plant cell walls, Ochiai *et al.*, 2010) and LysM domains (CAZymes known to overcome the chitinase defence response in plants, de Sain & Rep, 2015). In additon, there were other highly expressed CAZymes shared between isolates, which have been shown to be universally important in fungi to break down plant cell walls upon infection (Levasseur *et al.*, 2013), allowing the fungus to penetrate plant tissue and reach the xylem where it can spread systemically.

Highly expressed genes within 2 kb of a mimp and predicted to produce secreted proteins which were unique to FOP (such as g17046 in FOP1 EMR; g18880, g18891 and g19747 in F81; and g17990, g18001 and g18184 in R2) were also predicted as putative effectors by EffectorP, and therefore would be candidates for putative effectors conferring host specificity on pea. There were also a small number of highly expressed genes (predicted to produce secreted proteins and within 2 kb of a mimp) which were unique to isolates of each of the three FOP races. These had no functional annotation from the genome analysis (Chapter 3) but would be candidates for race specific effector genes. Further study of all of these genes would be required to determine if they were unique to FOP or FOP races or whether homologous genes are identified in other fungi. This could be achieved by BLAST searching the gene sequences to look for similarities with other effectors (e.g. using a database such as NCBI).

Determining the importance of putative effectors in FOP during infection requires additional investigation, through the use of gene knockout and proteomics studies. Gene knockout studies are the most common approach utilised to elucidate the function of effectors and involves replacing the functional gene with an antibiotic selection marker usually by homologous recombination (Selin *et al.*, 2016). These methods have been used successfully to determine the importance of many *SIX* genes in FOL during infection, with examples including *SIX1* (Rep *et al.*, 2004), *SIX3* (Houterman *et al.*, 2009) and *SIX6* (Gawehns *et al.*, 2014). A knockout system was developed by Sorensen *et al.* (2014) for the efficient transformation of *F. avenaceum* and other filamentous fungi, using a USER-Brick (Uracil Specific Excision Reagent) vector system. A recent method being exploited is the clustered regularly interspaced short palindromic repeat (CRISPR-Cas9). This consists of using a Cas9 nuclease guided to a specific target site where it generates a double stranded break at the target site, which initiates a repair mechanism resulting in mutations, therefore disrupting the open reading frame of the gene (Nodvig *et al.*, 2015; Selin *et al.*, 2016). Proteomics analysis would help determine if putative effector genes produced functional secreted proteins. The xylem sap could be extracted from multiple infected peas and the proteins analysed by mass spectrometry to determine their amino acid and therefore nucleotide sequence. These could be compared to the candidate effector genes from bioinformatics studies to determine the presence of these proteins in the host. Similar studies have already been conducted in FOL infected tomato plants (Houterman *et al.*, 2007) and to determine resistance genes in FOP infected peas (Castillejo *et al.*, 2015). These techniques could be used to determine the importance of highly expressed *SIX* genes and other putative effectors in FOP, to determine their effect on pathogenicity in pea.

In conclusion, this study has been the first to examine the presence of *SIX* genes across multiple isolates representing different races of FOP and explored their expression during the early stages of infection through qPCR and RNAseq. In addition, novel highly expressed putative effectors have been identified in FOP, which may play a role in host specificity. Finally, highly expressed race specific effectors were identified which could explain race specificity to pea cultivars, and the evolution of FOP races. These results also provide new gene targets for rapid molecular identification of FOP races using *SIX* genes or other unique novel effectors.

5. Determining pathogenicity of *F. oxysporum* isolates on pea

5.1 Introduction

Currently, the only definitive method to distinguish between pathogenic and non-pathogenic *F. oxysporum* isolates on pea is to carry out pathogenicity tests (Kraft & Pfleger, 2001). Screening in greenhouses or growth chambers allows plants to be challenged at any time of the year, without interaction with other pathogens and under controlled environmental conditions of humidity, light and temperature (Infantino *et al.*, 2006). A variety of pathogenicity tests exist for testing the virulence of *Fusarium* isolates towards pea including solid inoculum assays (Chittem *et al.*, 2015), pipetting liquid inoculum onto pea seedling roots (Zitnick-Anderson *et al.*, 2018) and the widely used method of trimming roots and dipping in a conidial suspension (Kraft & Haglund, 1978). All these pathogenicity tests have been developed to elevate the effect of *Fusarium* diseases in pea, with some tests more suitable for general root rotting species, and others for specifically testing FOP wilt (Infantino *et al.*, 2006).

Pathogenicity tests designed to produce rapid results have made use of the ability to grow peas in test tubes. One of the early versions of a test tube pathogenicity assay was developed for screening peas for resistance to *F. solani* f. sp. *pisi* and involved surface sterilising seeds, germination on filter paper before suspension in 1 x 10^6 conidia ml^{-1} in 0.1% water agar for 14 days (Whalley, 1984; Kraft & Kaiser, 1993). An improvement to this test was made by placing the germinated pea seeds in Milcap (polypropylene fibre) plugs suspended 2 cm above 10 cm of nutrient solution at the bottom of the tube. The peas were grown until the roots were 3 cm long and then inoculated by replacing the nutrient solution with a spore suspension such that the pea roots were immersed in the inoculum (Dyer & Ingram, 1990). Similar iterations of this test were used to evaluate the pathogenicity of other root rot causing *Fusarium* species on pea, including *F. culmorum*, *F. poae* and *F. oxysporum*, by assessing the level of necrosis on the roots (Ivic, 2014). A tube assay has also been used to determine the pathogenicity of *F. redolens* and *F. graminearum* on pea, chickpea and durum wheat, by sowing germinated seeds into 25

mL of inoculated peat/vermiculite mix placed in sealed test tubes (Taheri *et al.*, 2011). In all cases this test allowed the discrimination of pathogenic and non-pathogenic isolates of different root rot causing *Fusarium* species on pea by measuring their ability to cause root discolouration and necrosis. As mentioned previously (Chapter 2) this test is a very artificial assay and can allow non-pathogenic isolates to cause small amounts of root browning, so it should be used in conjunction with other tests to fully determine pathogenicity.

One of the main inoculation methods for testing the pathogenicity of root rot causing *Fusarium* species (e.g. *F. solani*) on pea is a seed inoculation assay. In this method, liquid cultures were initiated using agar plugs of mycelium and conidia filtered out after 6 days growth using cheesecloth. Seeds were scarified before being inoculated by soaking in spore suspensions overnight at room temperature. Seeds were grown in perlite and scored at 20 dpi for lesion size and severity of root browning (Kraft & Kaiser, 1993; Grunwald *et al.*, 2003). Similar tests for inoculation of seeds before sowing have involved the surface sterilisation and pre-germination of seeds before soaking in a conidial suspension for 5 minutes. Roots were scored 14 days after planting using a 0-3 disease score rating scale where 0 = non-pathogenic and 3 = highly pathogenic (coalescing lesions >1 cm) (Feng *et al.*, 2010). Variations of this test are common, but the basic principle of inoculating seeds before sowing is a general practice for determining the pathogenicity of isolates that cause root rot, such as *F. solani* (Grunwald *et al.*, 2003; Ondrej *et al.*, 2008; Porter, 2010) and *F. avenaceum* (Feng *et al.*, 2010). This seed inoculation assay is also well established for testing the pathogenicity of *Fusarium oxysporum* in onion, where seeds are immersed in spore suspensions before sowing, to determine the impact of infection at the seedling stage (Ozer *et al.*, 2004; Dissanayake *et al.*, 2009; Taylor *et al.*, 2013). Similar assays have also been used to assess the pathogenicity of *Fusarium* in other hosts, including bakanae disease of rice (Amatulli *et al.*, 2012) and crown rot in wheat (Li *et al.*, 2008).

The universally accepted pathogenicity test for Fusarium wilt of pea is the root dip method, which involves inoculating seedlings in the third-fourth node stage by removing plants grown in perlite and dipping and pruning the root system whilst it is immersed in a conidial spore suspension. Symptoms of Fusarium wilt were then observed over the duration of the study, until inoculated controls were dead (varied by cultivar), which

consisted of stunting, yellowing, dying of lower leaves, downward curling of leaf margins and usually plant death (Kraft & Haglund, 1978). This test has been frequently used, often with only minor modifications to the protocol and is a reliable method of characterising *Fusarium* isolates that cause wilt. A revised method for inoculating multiple plants at once was developed which involved sowing up to six seeds in an individual cell of a potting tray, such that the roots formed a cohesive plug when plants were removed, instead of individual plants. This collection of seedlings were inoculated by removing the bottom third of the root plug and the whole mass submerged in inoculum for 3-5 secs before re-sowing (Haglund, 1989). Although this approach is commonly referred to, most recent assays revert back to inoculation of individual plants to test isolate pathogenicity (Bani *et al.*, 2012; Sharma, 2011; Merzoug *et al.*, 2014). There is some variation in this technique as to whether the roots are trimmed while submerged in the inoculum (Kraft & Haglund, 1978; Neumann & Xue, 2003) or prior to immersion (Haglund, 1989; Bani *et al.*, 2012). Disease symptoms are usually assessed as the percentage of leaves showing wilt symptoms on each plant out of the total number of leaves, and in some cases assigning a visual severity score to each wilted leaf between 1 (healthy) and 5 (dead) (Bani *et al.*, 2012).

The root trimming and dipping method has been shown to be important when determining pathogenic FOP races. Races can be distinguished by the reaction of pea differential cultivars following inoculation. Cultivars will either be resistant (no observable disease) or susceptible (dead or severely stunted, wilted plants) when inoculated using the root trimming and dipping method (Kraft, 1994). This assay and the differential cultivars were used to identify the re-occurrence of FOP race 1 in Washington state after previously being dormant for 30 years while being controlled with resistant cultivars. It was distinguished from race 5 based on three cultivars: Little Marvel (susceptible to races 1 and 5); DSP (resistant to race 1 and susceptible to race 5) and New Era (resistant to race 1 and susceptible to race 5) (Kraft *et al.*, 1974). These researchers were also instrumental in identifying a new pathogenic race 6, which killed cultivars resistant to FOP races 1, 2 and 5 (Haglund & Kraft, 1979). Currently the only way to distinguish FOP races is through the use of the pea differential cultivars and the root dip pathogenicity test (Kraft & Pfleger, 2001).

Fusarium wilt is also a major problem in many other plant hosts, and hence the root dip method is a well-established and commonly used assay for identifying pathogenic isolates of other *F. oxysporum formae speciales*. *F. oxysporum* f. sp. *lycopersici* (FOL), the causal agent of Fusarium wilt in tomato, is one of the most devasting diseases in tomato, causing severe yield reduction in those grown under greenhouse and field conditions (Nirmaladevi *et al.*, 2016). The standard protocol of testing the pathogenicity of FOL isolates is to use a root trimming and dipping method very similar to the assay used for pea (Nirmaladevi *et al.*, 2016). Pathogenicity of the wilt causing pathogen in lettuce, *F. oxysporum* f. sp. *lactucae*, is also tested with a similar root trimming and dipping method (Garibaldi *et al.*, 2004). The main difference to the root dip test used for pea is that plants of tomato and lettuce are usually grown in soil mixture, instead of perlite or vermiculite.

In previous chapters, isolates of *F. oxysporum* causing root rot and wilt were determined based on *TEF* phylogeny and the presence/absence of *SIX* genes. Isolates from diseased peas from UK fields were grouped into distinct separate clades from known FOP isolates and also contained few or no *SIX* genes. In comparison, FOP isolates separated into clades based on their preliminary race type using *TEF* sequencing, and also displayed a race specific complement of *SIX* genes (race 1 = *SIX1*, *SIX6*, *SIX7*, *SIX9*, *SIX10*, *SIX11*, *SIX12* and *SIX14*, race 2 = *SIX1*, *SIX6*, *SIX9*, *SIX13*, *SIX14*, race 5 = *SIX1*, *SIX6*, *SIX13*). Pathogenicity tests were needed to confirm the difference between *F. oxysporum* root rot and FOP isolates and the race type of FOP isolates.

The main aim of this chapter was to determine the pathogenicity of *F. oxysporum* isolates using the three pathogenicity tests assessed in Chapter 2, and to determine the race type of isolates of *F. oxysporum* f. sp. *pisi* (FOP).

The specific objectives were:

1. To determine which *F. oxysporum* isolates cause root rot in pea using test tube and seed inoculation pathogenicity tests
2. To identify pathogenic FOP isolates causing wilt symptoms in pea using the root dip pathogenicity test
3. To identify the FOP race type of pathogenic isolates using the root dip pathogenicity test and pea differential cultivars.

5.2 Materials and Methods

5.2.1 Test tube pathogenicity assay

Twenty-two *F. oxysporum* isolates (PG15, 16, 18, 21, 57, 58, 59, 60, 61, 62, 63, 65, 72, 73, 74, 76, 77, 79, 85, FOP1, 2 and 5) and the non-pathogenic isolate Fo47 isolate were assessed for their pathogenicity using the test tube assay as described in Chapter 2 (Section 2.2.3.1). The isolates were chosen based on the different clades they belonged to following phylogenetic analysis of *TEF* sequences (Section 2.3.2). Pea seedlings (cv. Avola, Kings Seeds, Colchester) were inoculated using the test tube assay previously described in Section 2.2.3.2 (including non-inoculated control tubes) and symptoms scored based on root discolouration (Section 2.2.3.2, Figure 2.1) every 2-4 days until 16 dpi. Ten tubes per isolate (including a non-inoculated control) were arranged in racks in an alpha design (Genstat release 18.1, VSN international Ltd), to account for any variation in temperature within the incubator.

5.2.2 Seed inoculation pathogenicity test

As with the test tube pathogenicity assay (above), 23 isolates were used to inoculate pea seeds for two hours in a spore suspension before sowing in compost, following the method described in Section 2.2.3.1. Seeds (cv. Avola, Kings Seeds, Colchester) were sown in modular trays and arranged in an alpha design (Genstat), with four replicate trays per isolate, and one non-inoculated control treatment per block. Plants were maintained in a temperature-controlled glasshouse (23˚C day, 18˚C night, 16 h photoperiod) for 28 days until control plants began to senesce after flowering and pod setting. Seed germination and subsequent death of plants was recorded every 3 to 7 days. A repeat experiment was performed, but in this case, plants were monitored for 34 days as this was the point that the control treatment started to senesce. At the end of both experiments, plants were collected and roots were washed after being separated from the stems. Roots and shoots were pooled for each isolate in each replicate, weighed, then dried at 80˚C for 3-4 days and re-weighed.

5.2.3 Root dip pathogenicity tests

5.2.3.1 Testing the pathogenicity of FOP isolates

The root dipping and trimming method (Section 2.2.3.3) was used to test the ability of 17 *F. oxysporum* isolates to cause wilt on a susceptible cultivar of pea (cv. Little Marvel, D.T Brown, Suffolk). The isolates were chosen based on their positions in the *TEF* maximum likelihood tree (Section 2.3.2) and their complement of *SIX* genes (Chapter 4, Table 5.1). Four isolates of each preliminary FOP race, designated by the presence/absence of *SIX* genes (race 1 = *SIX1*, *SIX6*, *SIX7*, *SIX9*, *SIX10*, *SIX11*, *SIX12* and *SIX14*, race 2 = *SIX1*, *SIX6*, *SIX9*, *SIX13*, *SIX14*, race 5 = *SIX1*, *SIX6*, *SIX13*) , along with one isolate with no *SIX* genes (previously categorised as race 2) and four isolates from UK fields were included in the test (Table 5.1). The non-pathogenic isolate Fo47, and non-inoculated controls were also included. Seedlings were inoculated as previously described (Section 2.3.3.3) and arranged in a randomised complete block design (Genstat) consisting of 18 blocks, with each block containing one pot of every treatment, and three pots of non-inoculated control plants. Plants were maintained in a temperature-controlled glasshouse (23˚C day, 18˚C night, 16 h photoperiod) for 41 days, and the total number of leaves per plant as well as the number of wilted leaves per plant was recorded twice per week. At the end the experiment, plants were collected and roots were washed after being separated from the stems. Roots and stems were pooled for each isolate across all 18 blocks, weighed, then dried at 80˚C for 3-4 days and re-weighed.

Table 5.1 *Fusarium oxysporum* isolates selected for pathogenicity testing using the root dip test based on preliminary race typing, *SIX* gene presence/absence (chapter 4) and their position in the *TEF* phylogenetic tree (chapter 2).

Race	Isolate	Clade on *TEF* tree	*SIX* gene presence/absence								
			1	*6*[1]	*6*[2]	*7*	*9*	*10*	*11*	*12*	*13*	*14*
1	F79	Clade 6 (FOP race 1)	■	■	■	■	■	■	■	■		■
	FOP1 EMR		■	■	■	■	■	■	■	■		■
	FOP1		■	■	■	■	■	■	■	■		■
	CBS170		■	■	■	■	■	■	■	■		■
2	F81	Clade 7 (FOP race 2)	■		■		■				■	■
	F231		■		■		■				■	■
	F31		■		■		■				■	■
	FOP2		■		■		■				■	■
5	R2	Clade 3 (contains FOP5)	■		■						■	
	F40		■		■						■	
	PDA3b		■		■						■	
	FOP5		■		■						■	
Root rot	Fw-08-04	Clade 5	■		■							
	PG247	Clade 5										■
	PG2	Clade 3 (contains FOP5)										
	PG18	Clade 1 (Root rot)										
	PG3	Clade 1 (Root rot)			■							
Non-path	Fo47	Clade 4										

5.2.3.2 Comparing susceptibility of Avola and Little Marvel inoculated with *F. oxysporum* isolates

The root dipping method described above was used to test the reaction of two pea cultivars (Avola and Little Marvel) to different isolates of *F. oxysporum*, selected based on previous pathogenicity data (5.3.3.1) and *SIX* gene complements (Chapter 4). Two isolates of each preliminary race (previously assigned by *TEF* phylogeny (Chapter 2) and *SIX* gene complement: race 1 = *SIX1*, *SIX6*, *SIX7*, *SIX9*, *SIX10*, *SIX11*, *SIX12* and *SIX14*, race 2 = *SIX1*, *SIX6*, *SIX9*, *SIX13*, *SIX14*, race 5 = *SIX1*, *SIX6*, *SIX13*): FOP1 and FOP1 EMR (race 1), FOP2 and F81 (race 2), and FOP5 and R2 (race 5) and one isolate shown to cause severe root rot (PG18, Chapter 2) were included in the experiment with the non-pathogenic isolate Fo47 and a non-inoculated control. Seedlings were arranged in a randomised complete block design (Genstat) with each of the 16 blocks containing one replicate of each isolate/treatment combination. Plants were maintained in a temperature-controlled glasshouse (conditions as before), for 48 days, then collected, weighed and

dried at the end of the experiment in the same way. The experiment was repeated once, and the data from both experiments combined in the analysis.

5.2.3.3 Determining race type of FOP isolates using pea differential cultivars

Pea differential cultivars were inoculated using the root dipping method described above and in Section 2.3.3.3 to determine the race of different FOP isolates. Two isolates of each preliminary race (1, 2 and 5) based on *TEF* phylogeny (Chapter 2) and *SIX* gene complement (as above), previously shown to be pathogenic against the susceptible cultivar Little Marvel (5.3.3.1), were used to inoculate the four differential pea cultivars Darkskin Perfection (DSP), Mini, Sundance II and Little Marvel (as the susceptible control) (Table 5.5). Seeds were obtained from the Germplasm Resources Unit (GRU) at John Innes Centre, Norfolk, and initially bulked up by growing plants until maturity before drying and storing the seeds. Seeds were inoculated using the previously described root dip method, with the non-inoculated seeds treated with SDW only. Plants were arranged in a temperature-controlled glasshouse (conditions as before) in a randomised complete block design with each of the 14 blocks containing one plant of each treatment/cultivar. Plants were monitored for wilt symptoms for 41 days then collected and dried at the end of the experiment as before.

5.2.4 Statistical analyses

All statistical analyses were carried out in Genstat® (release 18.1, VSN international Ltd), with advice and support from Andrew Mead, Rothamsted Research. For the test tube pathogenicity assay, there was very weak evidence of small-scale variation (accounted for by the alpha design) shown by running a REML (Restricted or residual maximum likelihood) linear mixed model analysis, so it was then analysed using a simpler randomised block design with an ANOVA. Interpretation of this analysis was carried out by plotting ANOVA means and comparing significant results with the LSD (5%).

For the seed inoculation pathogenicity test, the proportion of germinated seeds out of the total sown was calculated and transformed using a logit transformation with an offset of one to account for zero or 100 values. As before, an alpha design was used to arrange the trays to account for small scale spatial variation, but after testing with a REML linear mixed model analysis there was very little evidence for this, so the data was analysed using a simpler randomised block design with an ANOVA. Means from the ANOVA (significant result) were compared with the 5% LSD value and back transformed means were plotted. The plant survival data for this test were analysed using a Generalised Linear Model (GLM) with logistic regression, with fitted terms of replicate (block) + treatment (isolate) and a dispersion parameter fixed at 1. Interpretations from the GLM were made by comparing t probabilities calculated with reference to the untreated control for each isolate, and by plotting the estimated mean proportions predicted from the regression model and their associated standard errors. The last timepoint for each replicate was combined to account for the differences in plant maturity, and these data were used to give an overall representation of the two replicates before they were collected. Each replicate was also analysed separately at each time point to track disease progression over time. Root and shoot fresh/dry weights were also analysed using an ANOVA, after no spatial variation was observed from the alpha design.

For all the root dip pathogenicity tests, the number of wilted leaves was calculated as a proportion of the total number and transformed using a logit transformation with an offset of one. The transformed data was analysed with an ANOVA, considering the randomised block design used in the experiments. Treatment means from significant ANOVA's were compared to the control and each other using the LSD at the 5% level. If any pea cultivars were collected at different times due to plant maturity then a last time point for each cultivar was used as previously, and then cultivars were also analysed separately to evaluate disease progression over time. Logit transformed ANOVA means of the proportion of wilted leaves were back transformed, considering the offset (by using the mean number of leaves from all plants for each isolate) and plotted onto graphs as the percentage of wilted leaves. The Pearson product-moment correlation coefficient between root and shoot fresh/dry plant weights and percentage wilt for each isolate was also calculated to determine if there was an association between final plant weight and percentage wilt.

5.3 Results

5.3.1 Test tube pathogenicity assay

The pathogenicity of 23 *F. oxysporum* isolates was tested by visually inspecting root browning ranked between a score of 0 and 6 on a susceptible pea cultivar in the test tube pathogenicity assay. An ANOVA revealed there were overall significant differences in symptom score between the isolates tested at all time points ($p < 0.001$). Average symptom scores (Section 2.3.3.2, Figure 2.1) at 16 dpi indicated that all isolates caused significantly more root browning compared to the non-inoculated control (LSD = 0.746, 5% level), with no visible symptoms seen on control roots (Figure 5.1 a and c). The most pathogenic isolates were PG60, PG16, PG18, PG61 and PG62 with mean scores ranging from 5.3 to 5.5, with the least pathogenic isolates being FOP1, PG58 and PG76 with scores of 3 or lower. The least pathogenic isolates caused significantly less root discolouration compared to all the other isolates (Figure 5.1 a). The non-pathogenic isolate Fo47 caused significant browning compared to the control, as did the least pathogenic isolate PG76, with an average score of 3.8 at 16 dpi.

Selected *F. oxysporum* isolates were examined for their effect on root browning over the time course of the experiment. Two highly pathogenic isolates (PG18 and PG60), two less pathogenic isolates (PG21 and PG79) and the least pathogenic isolate PG76 were compared to the non-pathogenic isolate Fo47 and the non-inoculated control treatment. Disease scores for all isolates increased over time (apart from the control) but clear differences in the rate of disease progression were observed (Figure 5.1 b). The least pathogenic isolate PG76 had a relatively slow rate of disease progression, with scores ranging from 0.8 (5 dpi) to 2.7 (16 dpi), while a higher rate of disease progression was observed in the more pathogenic isolates PG18 and PG60, with scores ranging from 2 and 1.9 (5 dpi) to 5.4 and 5.5 (16 dpi) respectively. There was a significant change in the average score for PG60 between 12 dpi and 14 dpi, from 3.4 to 4.8. The other less pathogenic isolates displayed slower rates of disease progression over time, from scores of approx. 2 (5 dpi) to over 4 (16 dpi).

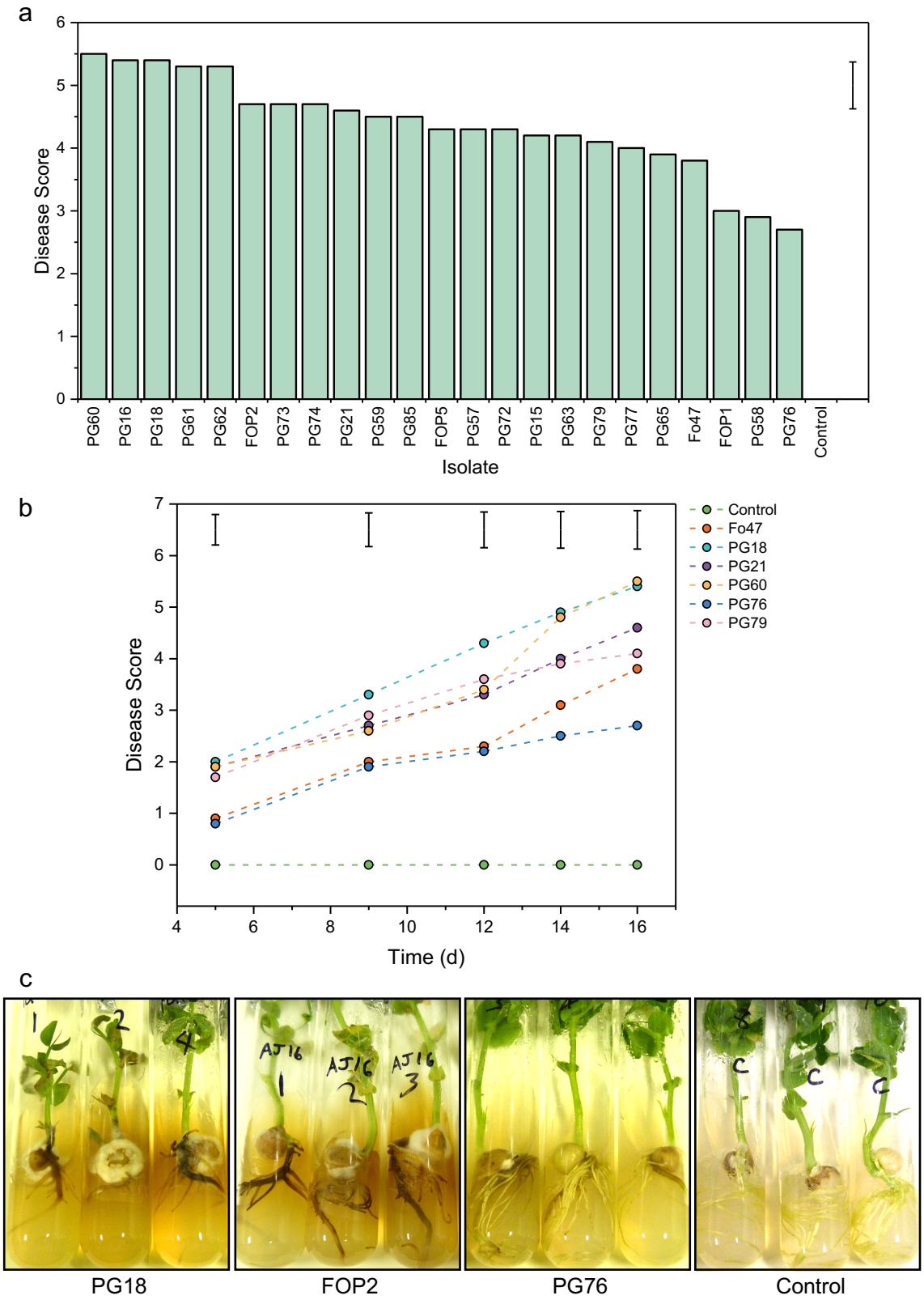


Figure 5.1 (A) Average disease score following inoculation of pea seedlings with 23 *Fusarium oxysporum* isolates using a test tube pathogenicity assay. Results show the average disease score at the final time point 16 dpi. (b) Average disease scores over time for a selection of six isolates from (a), scored at 5, 9, 12, 14 and 16 dpi. Error bars on both charts represent the least significant difference (LSD) at the 5% level. (c) Visual representation of root discolouration at 16 dpi for isolates: PG18, FOP2, PG76 and the Control.

5.3.2 Seed inoculation pathogenicity test

The percentage germination of pea seeds was measured to determine the level of pre-emergence damping off caused by each *F. oxysporum* isolate. As described in Section 2.2.3.1 the last time point from each experiment was combined to account for the slower maturity in the second experiment, and this combined data was used in the analyses. ANOVA revealed that there were significant differences between isolate means (including the control, $p < 0.001$). Comparing transformed means with a 5% LSD showed that all isolates caused a significant decrease in seed germination compared to the non-inoculated control (Table 5.2). The *F. oxysporum* isolate with the lowest percentage germination was PG62 at 57%, closely followed by PG18 at 62%, compared to Fo47 and PG65 with 87% of seeds germinated. No single isolate had a significantly greater effect on germination compared to the other isolates and instead the effect on germination is a continuum across the isolates tested (Figure 5.2a).

The percentage survival data of pea plants in the seed inoculation test were analysed using a General Linear Model (GLM) with logistic regression, using the combined data for the last time point. A significant difference in plant survival between *F. oxysporum* isolates was observed ($p < 0.001$, deviance ratio 9.05). All treatments, other than Fo47, caused a significant decrease in plant survival (of those that had germinated), compared to the control plants using a 5% LSD (Table 5.2). This can be interpreted using the back transformed predicted means to indicate that PG18 resulted in the lowest survival at 57%, compared to FOP1 at nearly 96% (Figure 5.2b). For all other isolates there was a continuum of survival, with only PG16 and PG18 showing significantly less survival compared to all other isolates (Figure 5.2b). Both FOP5 and FOP2 (historic isolates Warwick HRI) showed relatively low survival, with considerable reductions compared to Fo47 and FOP1, but still resulted in 81% and 77% of germinated plants surviving (Figure 5.2b, Table 5.2). Visual differences in plant survival and germination between the control and pathogenic isolates were observed throughout (Figure 5.2c).

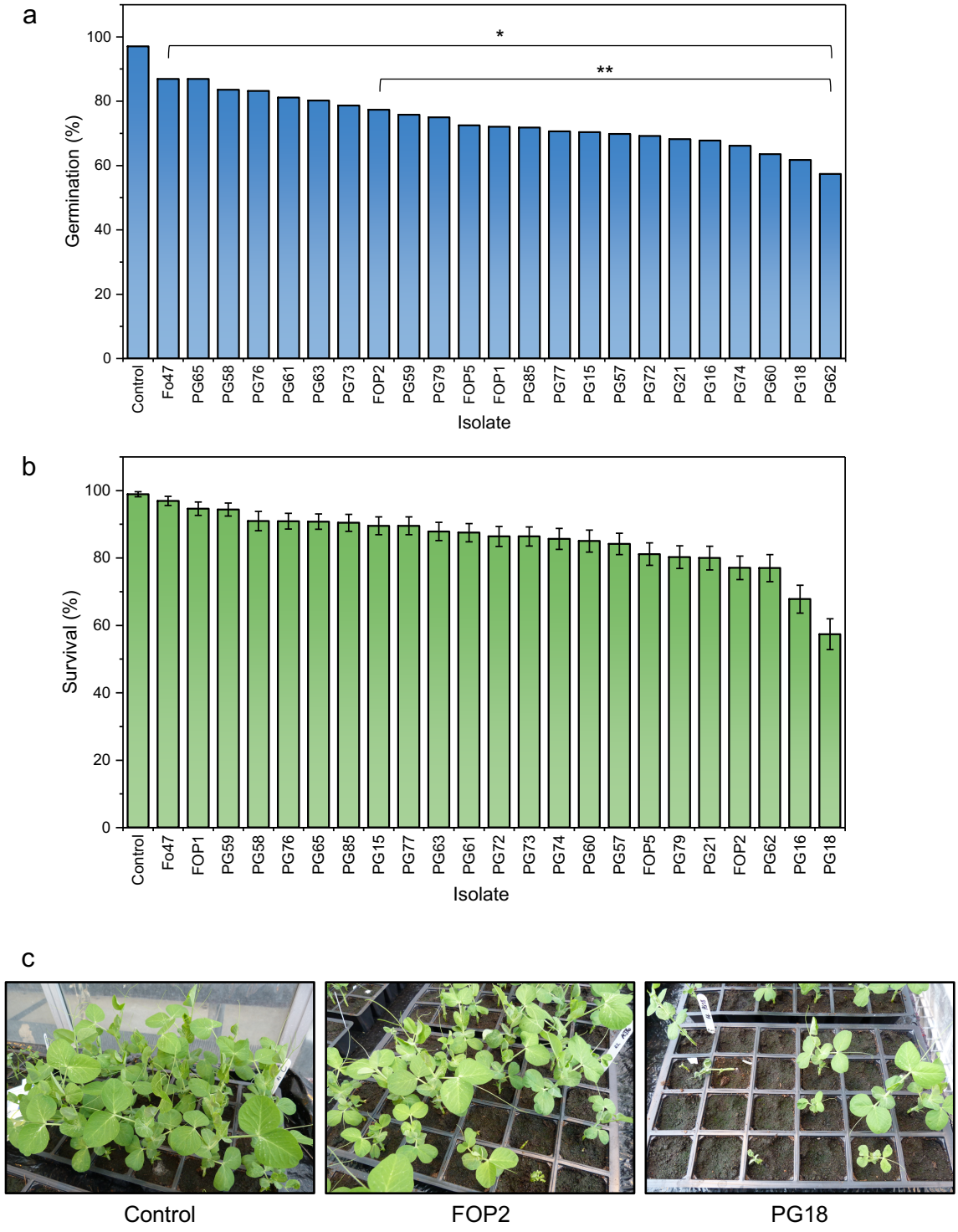


Figure 5.2 Pathogenicity of 23 *Fusarium oxysporum* isolates on pea using the seed inoculation method. (a) Data are back transformed means of percentage germination out of a total of 24 seeds sown, at the final time point across two replicates of the experiment. * all isolates caused significantly less germination compared to the control, ** all isolates caused significantly less germination compared to non-pathogenic isolate Fo47. (b) Survival (back transformed predicted means) of seedlings as a percentage of those that germinated, measured at the final time point across two replicates. Error bars represent predicted standard errors (from the GLM with logistic regression analysis). (c) Visual representation of germination and survival at 14 dpi for Control, FOP2 and PG18.

Table 5.2 Proportion of pea seeds germinated, and surviving plants for 23 *Fusarium oxysporum* isolates at the last time point of assessment from two replicates, using the seed inoculation pathogenicity test. Values represent transformed means from an ANOVA using a logit transformation with an offset (germination) and a GLM with logistic regression (survival). 5% LSD value used to compare significant differences between isolate means from the ANOVA. Higher values indicate higher % germination or high % survival.

Isolate	ANOVA (germination)	General linear model with logistic regression (survival)			
	Transformed means	Estimate	S.E.	t statistic	t probability
Control	2.66				
Fo47	1.66	-1.06	0.82	-1.28	0.199
PG65	1.66	-1.64	0.79	-2.08	0.037
PG58	1.45	-3.30	0.72	-4.61	<.001
PG76	1.43	-3.05	0.72	-4.24	<.001
PG61	1.31	-2.36	0.74	-3.18	0.001
PG63	1.26	-3.77	0.71	-5.29	<.001
PG73	1.18	-4.22	0.71	-5.92	<.001
FOP2	1.11	-3.13	0.72	-4.33	<.001
PG59	1.04	-2.84	0.73	-3.91	<.001
PG79	1.00	-2.20	0.77	-2.85	0.004
FOP5	0.88	-1.69	0.78	-2.17	0.03
FOP1	0.86	-2.77	0.73	-3.79	<.001
PG85	0.85	-2.56	0.73	-3.52	<.001
PG77	0.80	-3.30	0.72	-4.56	<.001
PG15	0.79	-2.53	0.73	-3.46	<.001
PG57	0.77	-2.22	0.74	-3.01	0.003
PG72	0.74	-2.66	0.73	-3.64	<.001
PG21	0.70	-2.66	0.73	-3.66	<.001
PG16	0.68	-2.72	0.73	-3.72	<.001
PG74	0.62	-2.21	0.74	-2.97	0.003
PG60	0.51	-2.36	0.74	-3.18	0.001
PG18	0.44	-3.11	0.72	-4.33	<.001
PG62	0.27	-2.26	0.75	-3.03	0.002
d.f	158				
5% LSD	0.53				

At the last time point for each seed inoculation experiment, roots and shoots were separated, weighed and dried before re-weighing. Increased isolate pathogenicity (Figure 5.2) was associated with a clear reduction in both the fresh and dry weight of shoots (Figure 5.3). All *F. oxysporum* isolates caused a significant reduction in weight (fresh and dry) of roots and shoots compared to the control (5% LSD, Figure 5.3) with the exception of Fo47 for fresh root, fresh shoot and dry shoot weights. However, in the dry root weight samples, Fo47 had a lower root weight (marginally significant) compared to control roots. As expected, the most pathogenic isolate, PG18 (Figure 5.2b), had the lowest plant/root weight in fresh and dry samples (Figure 5.3). Plants inoculated with the historic isolates FOP2 and FOP5 had low root and shoot weight, with significant reductions compared to the control and the non-pathogenic isolate Fo47, as well as other less pathogenic isolates.

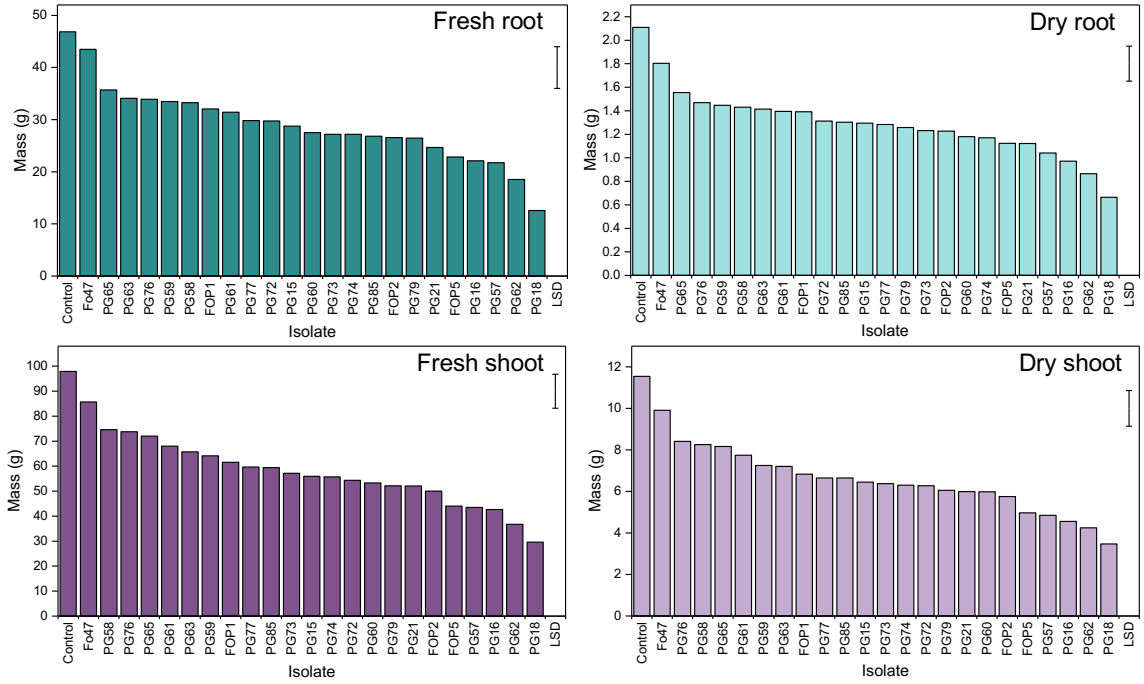


Figure 5.3 Weights of collected roots and shoots (fresh and dry) from pea plants inoculated with 23 *Fusarium oxysporum* isolates using the seed inoculation method. Error bars represent least significant difference (LSD, 5% level).

Two independent seed inoculation experiments were conducted but were collected at different time points due to differences in plant maturity. The percentage survival of plants over the course of the experiment was recorded, and a selection of isolates used to compare the progression of disease over time (Figure 5.4). Similar trends were observed in disease development between experiments, with most isolates either being consistently highly pathogenic (PG18), having moderate to low pathogenicity (PG21, PG60, PG76, PG79), or non-pathogenic (Fo47) (Figure 5.4).

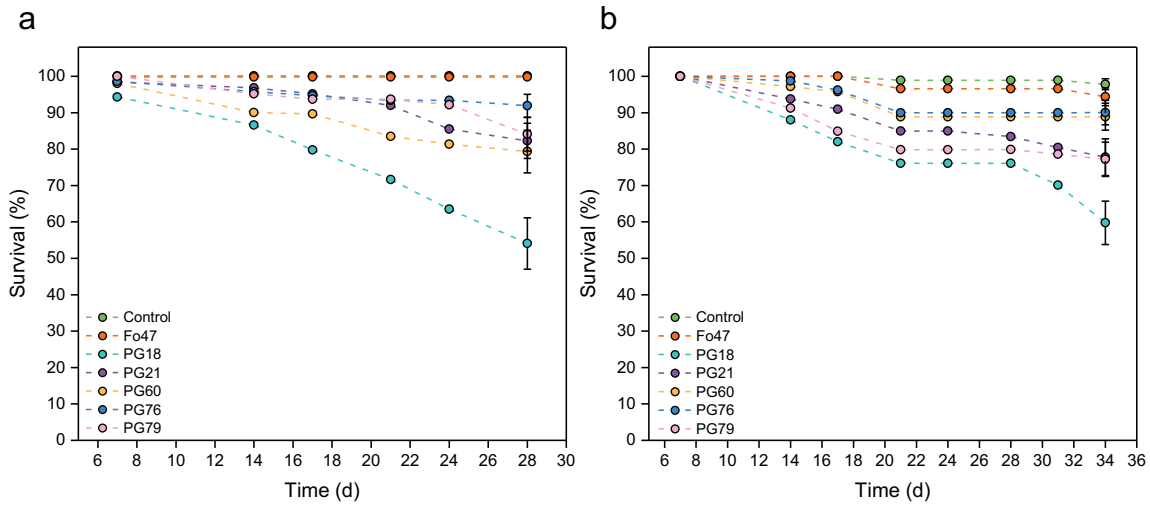


Figure 5.4 Percentage survival of pea plants (back transformed predicted means from GLM) out of the total number that germinated over time for selected *Fusarium oxysporum* isolates for two independent experiments using the seed inoculation method. (a) Results from the first experiment collected at 28 dpi, (b) results from the second experiment collected at 34 dpi. Error bars representing the predicted standard errors (predictions from the GLM with logistic regression analysis) are only shown for the last time point of each experiment, as this was used to compare isolate pathogenicity.

5.3.3 Root dip pathogenicity test

5.3.3.1 Testing pathogenicity of FOP isolates

The root dip pathogenicity test was used to screen selected FOP isolates for their ability to cause wilt on a universally susceptible pea cultivar (Little Marvel). Four isolates from each of the three putative races (1, 2 and 5), identified by *TEF* sequencing and *SIX* gene complement (Chapter 2, Chapter 4) were compared to isolates previously shown to cause root rot (5.3.2). The ANOVA using logit transformed data of the proportion of wilted leaves showed that overall there were significant differences between *F. oxysporum* isolates for the average proportion of wilted leaves ($p < 0.001$). These significant differences were observed when comparing isolate means to the control with a 5% LSD (Table 5.3), which showed that eight isolates caused highly significant levels of wilt. These included two race 1 isolates (F79 and FOP1 EMR), all the race 2 isolates and two of the race 5 isolates (R2 and F40), which all caused between 60% and 98% wilt (Figure 5.5a). There was greater variation in the amount of wilt caused by the race 2 isolates, with FOP2 causing significantly less wilt (60%) compared to F81 (85%) (Figure 5.5a, Table 5.3). None of the root rot causing isolates produced any wilt symptoms in this test, and hence were not significantly different from the uninoculated control (Figure 5.5a, Table 5.3). Two isolates each from races 1 (FOP1, CBS170) and 5 (FOP5, PDA3b) were found to be non-pathogenic, with no differences in percentage wilt compared to the control.

A selection of the most pathogenic *F. oxysporum* isolates was used to examine the rate of wilt development over time. The percentage of wilted leaves increased at each time point post inoculation from 13 to 41 dpi for each pathogenic isolate but the rate differed between isolates. Isolate R2 caused the fastest development of wilt symptoms with the greatest change observed between 16 and 20 dpi, where the percentage of wilted leaves increased by 55% from 14% to 69% in only four days. F79 also caused rapid wilt development but the rate was slower than with R2, as it took 30 days to cause more than 50% wilt (Figure 5.5b). Most other *F. oxysporum* isolates only began to cause an increase in disease development after 27 dpi, where percentage wilt started to increase linearly from around 20% at 27 dpi to over 85% at 41 dpi for FOP1 EMR, F81 and F40 (Figure 5.5b and c). FOP2 exhibited the slowest rate of disease development with 60% wilt at 41 dpi, but still caused a similar increase in rate after 27 days (Figure 5.5b).

All root and shoot fresh and dry weights displayed a moderate to strong negative correlation with average percentage wilt (Figure 5.6). However, isolates grouped into two distinct clusters, with pathogenic and non-pathogenic isolates being separated. Percentage wilt values of >60% resulted in a considerable reduction of root and shoot weights (both fresh and dry), and where there was little wilt, the root/shoot weights were higher. There was also an association between the severity of wilt and the reduction in plant weight (Figure 5.6).

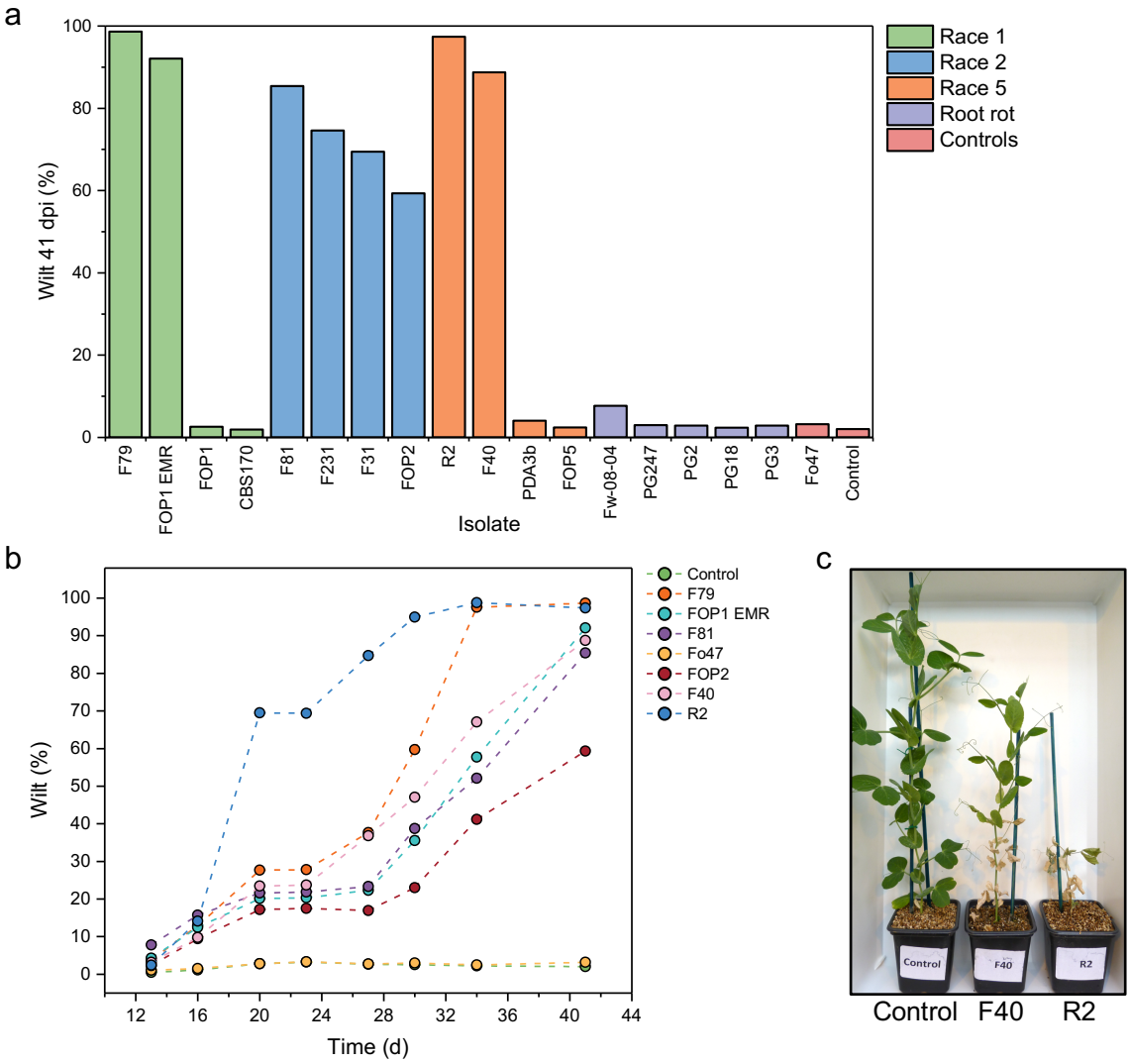


Figure 5.5 Pathogenicity of 18 *Fusarium oxysporum* isolates on a universally susceptible cultivar of pea (Little Marvel) measured as the average percentage leaves wilted, using the root dip inoculation method. (a) Percentage leaf wilt of all isolates at plant maturity, 41 dpi, with colours representing different races / groups of isolates included in the test. (b) Percentage wilt of a selection of pathogenic isolates and the non-pathogenic isolate Fo47 measured over time. All data are back-transformed mean values obtained following ANOVA analysis of logit transformed data. (c) Visual representation of pea plants inoculated with isolates F40 and R2 compared to control at 27 dpi.

Table 5.3 Pathogenicity of 18 *Fusarium oxysporum* isolates on a universally susceptible cultivar of pea (Little Marvel), 41 dpi, using the root dip pathogenicity test. Data represents logit transformed means for the number of wilted leaves as a proportion of the total number of leaves per pea plant following ANOVA analysis. 5% LSD values were used to compare significant differences between isolate means (Min.rep for comparing between all isolates other than control; Max.rep for comparing any isolate to the control). Higher numbers represent a greater proportion of wilted leaves.

Race	Isolate	Average proportion of wilted leaves (logit transformed)
1	F79	3.20
	FOP1 EMR	2.23
	FOP1	-3.22
	CBS170	-3.42
2	F81	1.67
	F231	1.04
	F31	0.79
	FOP2	0.36
5	R2	2.93
	F40	1.94
	PDA3b	-2.89
	FOP5	-3.25
Root rot	Fw-08-04	-2.34
	PG247	-3.11
	PG2	-3.15
	PG18	-3.29
	PG3	-3.16
Controls	Fo47	-3.07
	Control	-3.39
	d.f	341
Min.rep	5% LSD	1.13
Max.rep	5% LSD	0.65

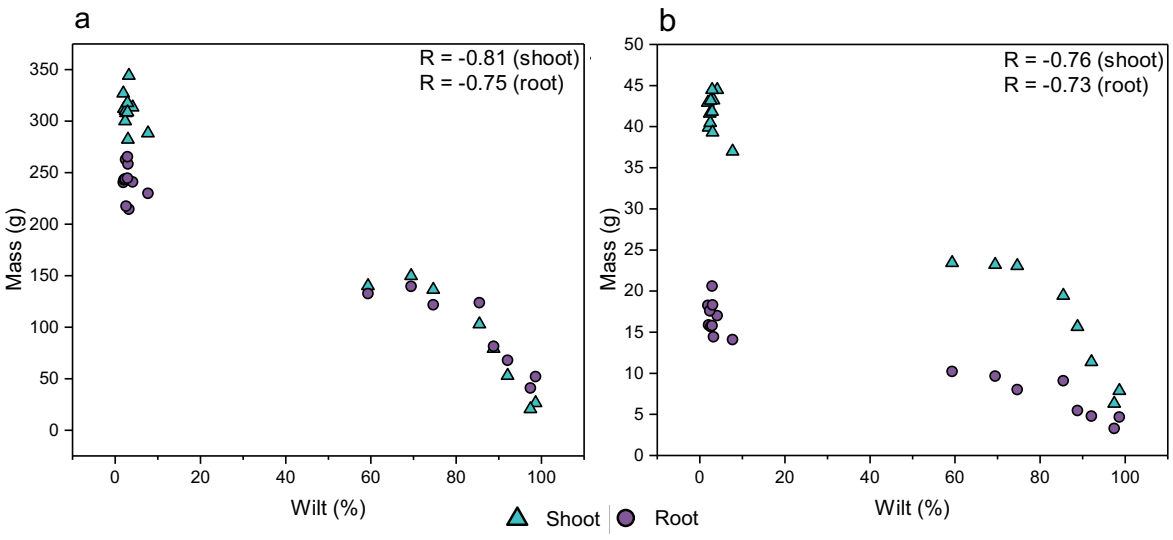


Figure 5.6 Correlation of pea fresh (a) and dry (b) root and shoot weights against percentage wilt for 18 *Fusarium oxysporum* isolates using the root dip inoculation test. R values represent the correlation coefficients.

5.3.3.2 Comparing susceptibility of Avola and Little Marvel inoculated with *F. oxysporum* isolates

Two cultivars of pea (Avola and Little Marvel) were inoculated with pathogenic and non-pathogenic *F. oxysporum* isolates from previous assays (seed inoculation and root dip pathogenicity tests) to determine the susceptibility of Avola to FOP race 1. Due to differences in maturity rate between the two cultivars tested, the last time point for each (28 dpi Avola, 48 dpi Little Marvel) was combined for the analysis so that comparisons between isolate and cultivar could be made at the same physiological time. The ANOVA revealed there was a significant interaction between *F. oxysporum* isolates across the two cultivars tested ($p < 0.001$). Comparing transformed isolate means of the proportion of wilted leaves using the 5% LSD (Table 5.4) showed that for Little Marvel, FOP1 EMR, FOP2, F81 and R2 all caused a significant amount of wilt compared to the uninoculated control. As before, the historic isolates FOP1 and FOP5 did not cause any significant wilt, with less than 6% of leaves wilted (Figure 5.7). All four pathogenic *F. oxysporum* isolates caused over 92% wilt in Little Marvel, supporting the results obtained previously (5.3.3.1). This was also the case for Avola, with the exception of FOP1 EMR which was non-pathogenic and caused no significant wilt (12%) compared to the control. PG18 and Fo47 had no significant effect on percentage wilt for both cultivars.

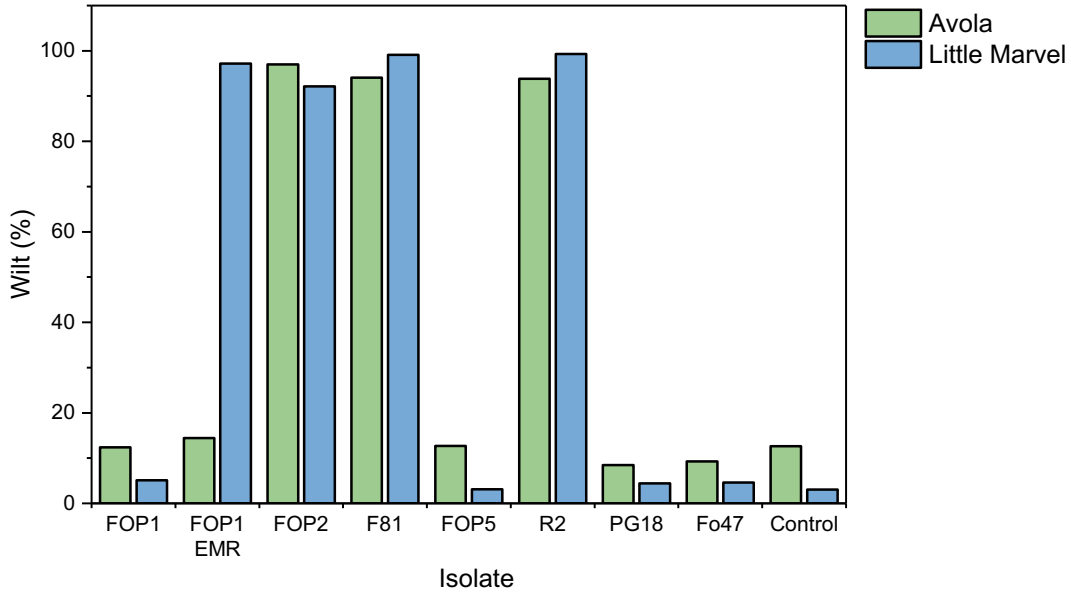


Figure 5.7 Pathogenicity of eight *Fusarium oxysporum* isolates on two cultivars of pea (Avola and Little Marvel) using the root dip inoculation method. Values for percentage wilt are back-transformed mean values obtained following ANOVA analysis of logit transformed data for a combination of two experiments; the last time point for each cultivar was combined due to differences in the rate of plant maturity.

Table 5.4 Pathogenicity of eight *Fusarium oxysporum* isolates on two cultivars of pea (Avola and Little Marvel), at the final assessment time point for each cultivar, using the root dip pathogenicity test. Data represents logit transformed means for the number of wilted leaves as a proportion of the total number of leaves per pea plant following ANOVA analysis, combining two experimental replicates. 5% LSD values were used to compare significant differences between means. Higher numbers represent a greater proportion of wilted leaves.

Race	Isolate	Average proportion of wilted leaves (logit transformed)	
		Avola	Little Marvel
1	FOP1 EMR	-1.60	2.95
	FOP1	-1.80	-2.69
2	F81	2.36	3.35
	FOP2	2.80	2.29
5	R2	2.34	3.45
	FOP5	-1.77	-3.10
Controls	PG18	-2.14	-2.81
	Fo47	-2.05	-2.77
	Control	-1.77	-3.10
	d.f	527	
	5% LSD	0.61	

Root and shoot dry weights for both pea cultivars (Avola and Little Marvel) displayed a strong negative correlation compared to the average percentage wilt. However, isolates grouped into distinct clusters with high root/shoot weights associated with low percentage wilt values, and low root/shoot weights with high percentage wilt values (Figure 5.8). There were no isolates which caused mild pathogenicity resulting in moderate wilt and intermediate weights (Figure 5.8).

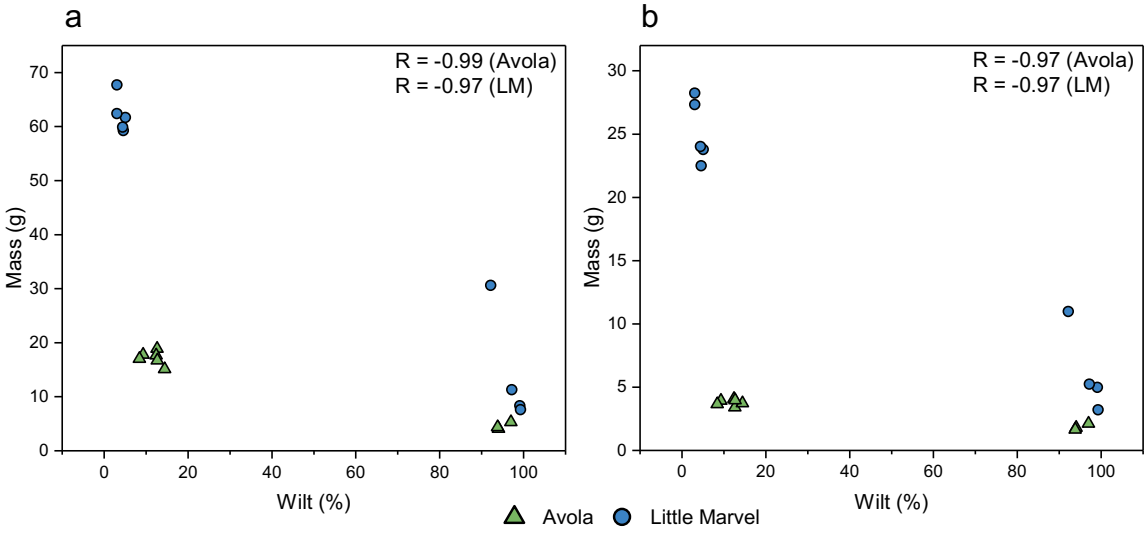


Figure 5.8 Correlation of pea shoot (a) and root (b) dry weights after collection against percentage wilt for Avola (green) and Little Marvel (blue) cultivars inoculated with eight *Fusarium oxysporum* isolates using the root dip inoculation test. R values represent the correlation coefficients.

Disease rate was observed over the time course of the experiment for each cultivar and showed that there were differences between pathogenic and non-pathogenic isolates for both cultivars. (Figure 5.9). In Avola, disease developed rapidly over the first 20 days with the three pathogenic isolates F81, R2 and FOP2, whereas all other isolates retain low levels of wilt across the time course. A similar trend can be observed with Little Marvel, with the four pathogenic isolates causing rapid wilt symptoms to appear over the first 35 days, with the non-pathogenic isolates resulting in low levels of wilt throughout. The same isolates were pathogenic between cultivars with the exception of FOP1 EMR, which was pathogenic on Little Marvel but not on Avola. FOP2 had a slower rate of disease development in Little Marvel compared to Avola, but still caused over 90% wilt by 48 dpi (Figure 5.9).

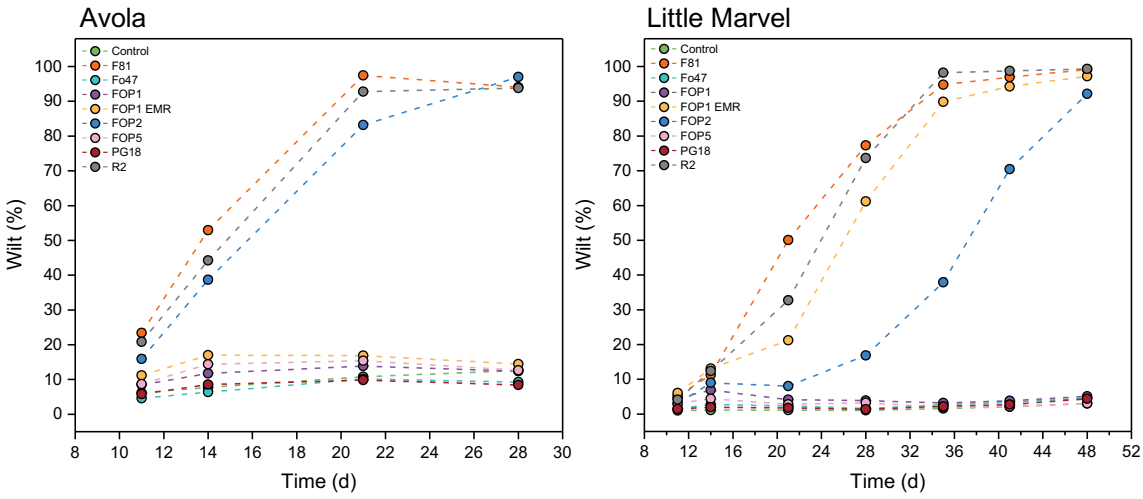


Figure 5.9 Percentage of wilted leaves (back transformed from an ANOVA using logit transformed data of the proportion of wilted leaves) over time for selected *Fusarium oxysporum* isolates used to inoculate two cultivars of pea (Avola and Little Marvel) as part of the root dip pathogenicity test. Cultivars were collected at different time points: Avola = 28 dpi; Little Marvel = 48 dpi.

5.3.3.3 Determining race type of FOP isolates using pea differential cultivars

The root dip inoculation method was used to inoculate four pea differential cultivars: Little Marvel, Mini, Sundance II and DSP with two *F. oxysporum* isolates from each preliminary race type (based on *TEF* phylogeny (Chapter 2) and *SIX* gene profiles (Chapter 4)); race 1 (FOP1 EMR and F79), race 2 (FOP2 and F81) and race 5 (R2 and

F40). All isolates tested were highly pathogenic on the susceptible cultivar Little Marvel as observed in previous tests (5.3.3.1, 5.3.3.2). ANOVA results revealed that there was a significant interaction between the isolates and cultivars tested ($p < 0.001$). For the universal susceptible cultivar Little Marvel, all isolates caused significant wilt (over 70% leaves affected) compared to the control using a 5% LSD (Table 5.6). DSP (a line resistant to FOP race 1 but susceptible to race 2 and 5) was resistant when inoculated with FOP1 EMR and F79 (race 1), but was highly susceptible to inoculation with FOP2, F81 (both race 2), R2 and F40 (both race 5) (Figure 5.10 and Figure 5.11). In addition, FOP1 EMR and F79 (race 1) caused highly significant wilt compared to the control in the cultivar Mini (susceptible to race 1 and 5, resistant to race 2) (Figure 5.10, Figure 5.11 and Table 5.6). This cultivar also displayed resistance to FOP2 and F81 (race 2) compared to the control (as expected). However, Mini was also highly resistant to *F. oxysporum* isolates R2 and F40, which had both been initially designated as race 5 based on *SIX* gene complement. Sundance II was highly resistant to all isolates, and no significant wilt was observed compared to the control. Results from this experiment were directly compared to the expected result of resistance or susceptibility of each cultivar to the different races (Table 5.5). Based on these pea differential tests, *F. oxysporum* isolates FOP1 EMR and F79 were confirmed as race 1. However, FOP2 and F81 (race 2), and R2 and F40 (race 5) could not be distinguished with the differential cultivars used, but based on the result of Mini and *SIX* gene profiles (Chapter 4), it is likely that they were all a subset of race 2.

Table 5.5 Summary of expected (Exp.) and observed (Obs.) results for the resistance and susceptibility of pea differential cultivars to different *Fusarium oxysporum* f. sp. *pisi* (FOP) races. Pea cultivars included were Little Marvel, DSP, Mini and Sundance II which were inoculated using the root dip test with isolates representing races 1, 2 and 5 as initially designated based on *SIX* gene complement (Table 5.1). S = susceptible, R = resistant.

Race	Isolate	Little Marvel		DSP		Mini		Sundance II	
		Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.
1	FOP1 EMR	S	S	R	R	S	S	R	R
	F79	S	S	R	R	S	S	R	R
2	FOP2	S	S	S	S	R	R	S	R
	F81	S	S	S	S	R	R	S	R
5	R2	S	S	S	S	S	R	R	R
	F40	S	S	S	S	S	R	R	R

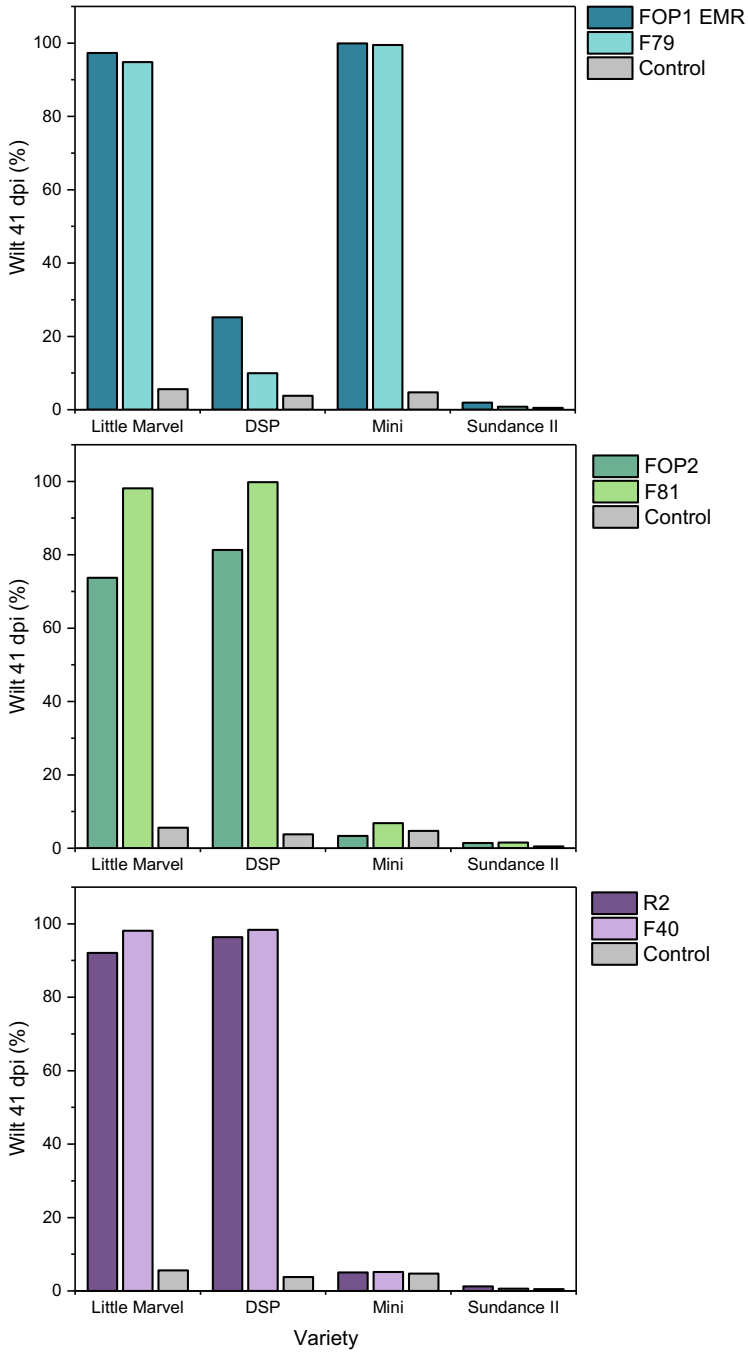


Figure 5.10 Pathogenicity of six *Fusarium oxysporum* f. sp. *pisi* (FOP) isolates on four differential cultivars of pea (Little Marvel, DSP, Mini and Sundance II) using the root dip inoculation method. Top = preliminary race 1 isolates; middle = preliminary race 2 isolates and bottom = preliminary race 5 isolates. Values for percentage wilt are back-transformed mean values obtained following ANOVA analysis of logit transformed data.

Table 5.6 Pathogenicity of six *Fusarium oxysporum* f. sp. *pisi* (FOP) isolates on four differential cultivars of pea (Little Marvel, DSP, Mini and Sundance II) 41 dpi. Data were transformed means for the number of wilted leaves as a proportion of the total number of leaves per pea plant following ANOVA analysis of logit transformed values for the root dip pathogenicity test.

Race	Isolate	Average proportion of wilted leaves (logit transformed)			
		Little Marvel	DSP	Mini	Sundance II
1	FOP1 EMR	2.75	-0.71	3.03	-3.26
	F79	2.42	-1.47	2.90	-3.65
2	FOP2	1.94	2.95	-2.41	-3.51
	F81	3.01	3.33	-0.90	-3.32
5	R2	2.55	3.27	-2.08	-3.53
	F40	2.93	3.34	0.11	-3.23
	Control	-2.49	-2.52	-2.71	-3.23
	d.f				348
	5% LSD				1.10

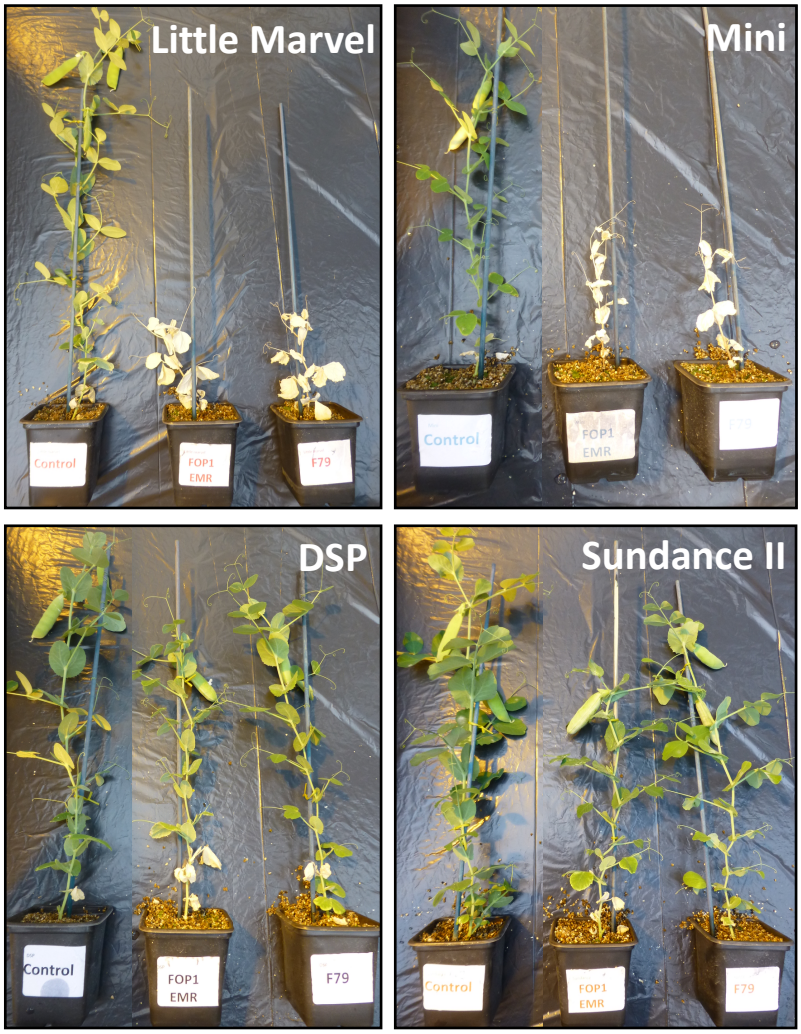


Figure 5.11 Visual representation of symptoms of *Fusarium oxysporum* f. sp. *pisi* (FOP) infection on the four differential pea cultivars from the (left to right) uninoculated control, FOP1 EMR (race 1) and F79 (race 1) treatments at 41 dpi.

5.4 Discussion

In this chapter, pathogenicity tests (Chapter 2) were used to differentiate between root rot *F. oxysporum* isolates from diseased UK peas, and FOP isolates which had been previously race typed. The seed inoculation test showed that *F. oxysporum* isolates from diseased peas in UK fields caused root rot symptoms in a susceptible pea cultivar, whereas they caused no wilt symptoms in the root dip pathogenicity test (for determining wilt), unlike FOP isolates which caused severe wilt. Race typing of putative FOP races, determined by *TEF* gene phylogeny and *SIX* gene screening (Chapter 2, Chapter 4), revealed that the putative race 5 isolates were possibly a variant of race 2, and therefore cannot be distinguished using *SIX* gene screening. FOP races 1 and 2, based on their different *SIX* gene complements (Chapter 4), were confirmed by the results from race typing. This study directly compares *F. oxysporum* as a root rot and wilt causing pathogen in pea through the use of pathogenicity tests, and to relate races of FOP with differing complements of *SIX* genes.

Variations of the test tube pathogenicity assay have been used previously to test the virulence of suspected root rot causing *Fusarium* species in pea (Dyer & Ingram, 1990; Ivic, 2014). In one of these studies, it was found that nearly all *Fusarium* isolates caused root necrosis, but they differed in severity, with *F. oxysporum* isolates having median scores of 2 and 2.5 (Ivic, 2014). In the current study with 23 *F. oxysporum* isolates, there was a range of pathogenicity scores recorded across the isolates tested. All isolates caused a significant amount of root browning compared to the control, and those resulting in the lowest scores (FOP1, PG58 and PG76) were significantly different from all other isolates. This range of pathogenicity is supported by the findings of Ivic (2014), but as the test used here relies on visual inspection of roots and a subjective score being assigned between 0 – 6, it is subject to large amounts of variation. The scoring system categories encompass large differences in percentage root area, for example 50 – 90%, which limits the ability of the test to finely dissect differences in isolate virulence. This test was also highly variable, with scores ranging from 1 to 5 for plants inoculated with Fo47, which was likely due to variation in the pea seeds, and when grown under stressful conditions, may have made them more susceptible to a normally non-pathogenic isolate. Nonetheless, the test tube pathogenicity assay was very rapid (only 16 days compared to up to 34 days

for the seed inoculation test), and allowed *F. oxysporum* isolates causing high levels of root rot to be identified quickly.

Seed inoculation assays have been used extensively to examine the pathogenicity of root rot causing *Fusarium* species in pea, such as *F. solani* (Grunwald *et al.*, 2003; Ondrej *et al.*, 2008) and *F. avenaceum* (Feng *et al.*, 2010). There has been fewer studies on the effect of *F. oxysporum* as a root rot pathogen, and although it is frequently isolated from diseased peas, it is not always the most pathogenic species (Chittem *et al.*, 2015). In the current study however, the seed inoculation test showed that all *F. oxysporum* isolates caused a significant decrease in the percentage of germinated pea seeds and their subsequent survival, compared to the control. These results confirm that *F. oxysporum* isolates collected from diseased peas in UK fields were causing root rot and not wilt, which was hypothesised due to them grouping into separate lineages in the *TEF* phylogenetic tree (Chapter 2). However, only three isolates (PG18, PG60 and PG62) resulted in less than 65% germination, which could be due to most isolates having less of an effect on pre-emergence damping off, or that seeds were not soaked for long enough as many previous studies soaked at least over night or for 24 - 48 hrs (Kraft & Kaiser, 1993; Ondrej *et al.*, 2008). As with the test tube pathogenicity assay, there was a continuum of virulence seen over the isolates tested with the seed inoculation test, and only isolates at extreme ends differed significantly. PG18 was confirmed to contain no *SIX* genes (Chapter 4) but resulted in high plant mortality, and caused the most browning in the test tube pathogenicity assay, suggesting that *SIX* genes are not necessary for pathogenicity in *F. oxysporum* isolates causing root rot.

Isolates of FOP were included in the seed inoculation pathogenicity test to see if they could also cause root rot as well as wilt. FOP2 (race 2) caused 100% wilt in preliminary tests (Chapter 2) and high levels of wilt in this study in the root dip pathogenicity test. However, in the seed inoculation test it caused no wilt symptoms on inoculated plants, and although it caused relatively low percentage survival compared to the control and the other isolates tested, 77% of plants (of those that germinated) survived until collected. The relatively low plant survival could be due opportunistic infection of plants being grown under reasonably stressful conditions. FOP1 (race 1) caused no decrease in plant survival compared to the non-pathogenic isolate Fo47, although it also showed no wilt symptoms in plants in the root dip pathogenicity test in this study and the preliminary test

(Chapter 2). Ondrej *et al.* (2008) reported similar results with the seed inoculation test as little root rot was caused by a FOP race 1 isolate, but a considerable reduction in seedling survival was observed for a race 2 isolate, similar to that of the *F. solani* isolates tested.

As mentioned previously, seed inoculation assays have frequently been used to examine the pathogenicity of other root rot *Fusarium* species (*F. solani* and *F. avenaceum*), with many studies using a visual scoring method for root discolouration, similar to that for the test tube pathogenicity assay (Kraft & Kaiser, 1993; Grunwald *et al.*, 2003). This could be an alternative way of scoring for root rot symptoms, although importantly it does not consider pre-emergence damping off and post-emergence damping off. Scoring could only take place once, at the point of plant collection, so disease progression over time would not be accounted for either. Therefore, the assay described here is more appropriate for assessing multiple symptoms of root rot.

The pea root dip pathogenicity test is regarded as the standard inoculation method for assessing FOP isolates for their ability to cause wilt in pea (Kraft & Haglund, 1978; Haglund, 1989). It has also been frequently used as a method of screening for resistance to FOP in pea cultivars (McPhee *et al.*, 1999; Bani *et al.*, 2012) and to determine pathogenic races of FOP (Haglund & Kraft, 1970; Haglund & Kraft, 1979). In this study, the assay was used to determine the pathogenicity of different FOP isolates, preliminarily race typed as race 1, 2 and 5 based on *SIX* gene complements and *TEF* phylogeny (Chapter 2, Chapter 4). The pea cultivar Little Marvel was used in this test as it is the universally susceptible cultivar in the documented set of pea differentials (Kraft, 1994; Kraft & Pfleger, 2001), while Avola, which was also used in several of the above tests, was found to be resistant to FOP race 1 (Seminis, 2013). This could explain the lack of pathogenicity of the historic FOP1 isolate in the seed inoculation pathogenicity test. However, FOP1 was also non-pathogenic towards Little Marvel in the root dip test. There was also some variation in pathogenicity of the race 2 isolates in the root dip test. One explanation of this could be that when the reclassification of FOP races occurred, many race 3 and 4 isolates were found to be more virulent cultures of race 2 (Kraft, 1994; Kraft & Haglund, 1978), and therefore virulence of isolates can vary. *F. oxysporum* isolates that caused root rot (PG247, PG2, PG18, PG3 and Fw-08-03) which clustered together in the *TEF* phylogeny and lacked *SIX* genes, were non-pathogenic in this assay, supporting the hypothesis they cause root rot but not wilt. FOP isolate R2 (one of the preliminary

race 5 isolates), caused the most rapid disease development over time, indicating that this was a highly aggressive isolate. Two of the historic FOP isolates (FOP1 and FOP5) from Warwick HRI were both non-pathogenic in this test, and due to the uncertainty of their origin it is not clear whether these belong to race 1 and 5 (respectively). This is unlikely as they have the same *SIX* gene profile as other pathogenic race 1 and 5 isolates, and is more likely to be due to loss of virulence genes. Studies have reported that repeated transfer of a pathogen on artificial media can result in loss of virulence caused by genetic changes driven by the lack of a selection pressure for virulence (a host) in a lab environment (Smith *et al.*, 2008). Spontaneous loss of LS chromosomes has also been observed in FOL, in addition to deletion hotspots in LS regions where loss of effectors could occur (Vlaardingerbroek *et al.*, 2016). In this case it is unlikely to be a whole LS chromosome due to the presence of *SIX* genes in these isolates, as these were shown to be in LS regions in FOP (Chapter 3).

Due to the differences in the pea cultivar used between the root dip pathogenicity test (Little Marvel) and the seed inoculation pathogenicity test (Avola), a replicated experiment was conducted with known pathogenic and non-pathogenic FOP isolates and *F. oxysporum* isolates causing severe root rot. Avola (susceptible to root rot but resistant to FOP race 1) (Seminis, 2013) and Little Marvel (the universally susceptible cultivar from the pea differential set) (Kraft, 1994; Kraft & Pfleger, 2001) were inoculated with the same set of isolates using the root dip pathogenicity test. FOP1 and FOP5 were non-pathogenic on both Avola and Little Marvel, supporting results from the previous root dip test. The root rot isolate PG18 was also non-pathogenic towards both cultivars using this method of inoculation confirming its inability to cause wilt despite high virulence towards Avola in the seed inoculation test. The lack of wilt in Avola when inoculated with FOP1 EMR supports the evidence that it is resistant to FOP race 1 (Seminis, 2013), as this isolate caused high percentage wilt on Little Marvel which is known to be susceptible (Kraft & Pfleger, 2001). FOP2 caused 100% wilt on Avola, confirming its identity as an aggressive vascular wilt pathogen. The rate of disease development over time was different in both pea cultivars when inoculated with FOP2, which could be due to variation in the susceptibility of cultivars to certain races.

Pathogenicity tests using pea differentials are currently the only means of determining the pathological classification and race types of FOP isolates (Grajal-martin *et al.*, 1993).

In this study, four of the pea differential cultivars were used to race type six FOP isolates. The universal susceptible Little Marvel was as expected susceptible to all isolates, supporting the previous tests carried out in this study. DSP displayed resistance to isolates FOP1 EMR and F79, also as expected, suggesting that both isolates belong to race 1. However, both of the cultivars Mini and Sundance II did not react as expected when challenged with isolates initially identified as race 2 or 5 based on *SIX* gene complements. Mini, as expected was susceptible to the race 1 isolates (FOP1 EMR and F79) and resistant to race 2 isolates (FOP2 and F81), suggesting they have both been correctly assigned to the correct race. However, Mini was also resistant to the putative race 5 isolates R2 and F40 suggesting that these were incorrectly assigned and are more likely a subset of race 2. Sundance II showed resistance to all the FOP isolates which was as expected for race 1 and race 5 but not for race 2. Therefore, the expected outcome for race 2 and 5 has been reversed in Sundance II and Mini respectively so it is not possible to distinguish these as either race 2 or 5. However, the reaction of differential cultivars to race 2 isolates has previously been reported to be variable (Kraft, 1994), therefore it is probably more likely that they are different varieties of race 2, than race 5. The reaction of New Season differential cultivar has been reported to vary with different isolates of race 2 (Kraft, 1994; Kraft & Pfleger, 2001), therefore this could also be the case with these isolates and cultivars. To resolve this, a further experiment would be required including additional pea differential cultivars New Era, WSU 28 and New Season (although potentially variable) to try and distinguish between race 2 and 5. Overall, it was only FOP1 EMR and F79 that could be confirmed as race 1 isolates as they caused the expected reaction in all four pea cultivars.

In conclusion, the use of different pathogenicity tests has enabled isolates causing root rot and wilt to be distinguished, and their pathogenicity to be determined. The test tube and the seed inoculation pathogenicity tests both displayed the effects of *F. oxysporum* isolates that cause root rot symptoms, such as blackening of the upper tap root and area of cotyledon attachment, and damping off of young seedlings. In both of these tests there were no wilt symptoms on leaves, only loss of root weight and discolouration. In contrast, the root dip pathogenicity test displayed symptoms of wilt caused by FOP isolates, and there was only mild discolouration on some of the roots. The oldest leaves wilt and dry out first, and this continues up the plant until it is completely dry and fully wilted. In some races this occurs before pod setting and therefore no pods or only empty pods are

produced (Kraft & Pfleger, 2001). Two isolates in particular (PG18 and FOP2) showed the differences between these tests and how they can be used to distinguish between root rot and wilt causing isolates. PG18 caused severe pre and post emergence damping off in the seed inoculation pathogenicity test, but in the root dip pathogenicity test it caused no wilt symptoms and plants developed healthily. In contrast, FOP2 caused far less of an effect on seedling emergence and survival in the seed inoculation test and caused no wilted leaves, but in the root dip pathogenicity test it caused severe wilt symptoms with almost 100% of the leaves affected in most tests.

Improvements in this study could include incorporating more of the same isolates in both pathogenicity tests to thoroughly assess the differences in isolate symptoms between assays. However, due to replication requirements and restrictions on experimental size based on glasshouse compartments, it was not possible in the current study, and further tests would have been needed. Also, additional cultivars could have been included when testing FOP isolates to determine their race type, but with the four cultivars and seven isolates (including the control) the maximum capacity of space and feasibility of experimental set up was quickly fulfilled.

6. General discussion

Fusarium wilt in pea has been reported in every country where peas are grown commercially, and has potential to cause severe crop losses. Although it is currently controlled by growing resistant cultivars, there is a constant risk of resistance break down due to the use of single gene sources of resistance being overcome by new pathogen races, as has occurred previously (Bani *et al.*, 2012). There have been very few studies evaluating the range of *Fusarium* species affecting UK peas, and relatively little is known about the pathogenicity factors in different races of *F. oxysporum* f. sp. *pisi* (FOP). Mechanisms of pathogenicity in *F. oxysporum* have been widely studied in numerous *formae speciales* (f. spp.), leading to the identification of effector genes shown to be vital for full pathogen virulence (de Sain & Rep, 2015).

The aim of this research (outlined in Section 1.12) was to determine the *Fusarium* species causing disease in UK peas and to understand the genetic basis for pathogenicity of different *F. oxysporum* f. sp. *pisi* (FOP) races.

The specific objectives were to:

1. Identify the different *Fusarium* species affecting peas in the UK and determine the importance of *F. oxysporum* as a pathogen.

2. Identify putative effector genes in three races of FOP using whole genome sequencing.

3. Evaluate the expression of putative effector genes in FOP races *in planta*.

4. Verify the race type of pathogenic FOP isolates using pathogenicity tests with pea differential cultivars.

Isolations from diseased peas from UK fields revealed a range of *Fusarium* species infecting pea roots and lower stems, with *F. oxysporum*, *F. solani* and *F. redolens* the most commonly isolated. Previously race-typed FOP isolates were also obtained mainly from the USA. Phylogenetic analysis based on *TEF* gene sequences from the *F. oxysporum* isolates (UK) as well as race-typed FOP isolates (USA), resulted in the *F.*

oxysporum isolates being grouped into two main clades, separated from the FOP isolates, which formed into clades based on their putative race. Therefore, it was hypothesised that isolates from diseased UK peas were associated with root rot and not FOP wilt. Other studies of diseased peas in North Dakota, Canada and the Netherlands also found a diverse range of *Fusarium* species causing root rot (Chittem *et al.*, 2015; Fernandez, 2007; Oyarzun *et al.*, 1993). Isolates were distinguished with the development of different pathogenicity tests, and to our knowledge this is the first direct comparison between *F. oxysporum* isolates causing root rot and wilt. There was inconsistency in the range and number of isolates used in the pathogenicity tests to distinguish *F. oxysporum* root and FOP isolates, and in future a wider selection of isolates should be included in each test. However, due to space and time constraints it was not possible to include a wider selection of isolates in this study. Nonetheless, the assays developed were appropriate for distinguishing between root rot and wilt symptoms. The absence of FOP isolates in diseased peas from UK fields could be due to the use of FOP resistant cultivars reducing incidence and therefore allowing root rot *Fusarium* species to establish. However, the method of isolation could have attributed to only root rot isolates being identified, as most samples were taken from the roots or lower stem and it is recommended to sample from the upper stem for isolating FOP (Kraft & Pfleger, 2001). Future work would include consistent sampling from multiple areas in the field to understand the spatial distribution and diversity of *Fusarium* species affecting individual fields or growing regions.

Genome sequencing was conducted for three isolates representing FOP races 1, 2 and 5 using long-read (MinION) technology. Comparative genomics revealed the distinction of the FOP genomes into core chromosomes and lineage specific (LS) regions, which conforms with other *F. oxysporum* f. spp. such as f. sp. *lycopersici* (FOL) (Ma *et al.*, 2010), f. sp. *radicis-cucumerinum* (van Dam *et al.*, 2017) and f. sp. *cepae* (Armitage *et al.*, 2018). LS regions revealed the presence of homologs of *Secreted In Xylem* (*SIX*) genes previously associated with pathogenicity in other *F. oxysporum* f. spp. (Houterman *et al.*, 2007; Schmidt *et al.*, 2013; de Sain & Rep, 2015). Significantly, there were different complements of *SIX* genes in each FOP race. The putative race 1 isolate (FOP1 EMR) contained *SIX1*, *SIX6*, *SIX7*, *SIX9*, *SIX10*, *SIX11*, *SIX12* and *SIX14*, whereas the putative race 2 isolate contained only *SIX1*, *SIX6*, *SIX9*, *SIX13* and *SIX14*, and the putative race 5 isolate contained *SIX1*, *SIX6* and *SIX13*. This result was confirmed in additional isolates of each race, and could therefore be responsible for host specific infection of FOP

in pea. In addition, the *SIX* genes were also found to be expressed *in planta* using qPCR, however, in most cases they were not the highest expressed genes when compared to the rest of the transcriptome using RNAseq analysis. The identification of highly expressed novel putative effectors was determined based on their location on LS regions, presence of a secretion signal and being located within 2 kb of a mimp, which are therefore likely to be involved in pathogenicity in FOP. A similar approach was used to determine novel effectors in FOL (Schmidt *et al.*, 2013).

SIX genes were mostly absent from *F. oxysporum* root rot isolates from UK fields and provided a method to distinguish them from FOP isolates using molecular methods. The presence of race specific *SIX* genes in FOP isolates, and the lack of them in root rot isolates, in addition to other FOP specific and race specific putative effectors identified by orthology analysis could provide targets for molecular diagnostics. This is useful for the agronomy sector in that fields containing wilted peas could be rapidly tested to determine if disease resistance had broken down (FOP) or if disease is due to root rot. Using the presence/absence of *SIX* genes to distinguish f. spp. and races has been determined previously (Lievens *et al.*, 2009), however this is the first report of a *SIX* gene race structure in FOP.

Furthermore, to truly understand the importance of the predicted putative effectors in this study, further work including gene knockouts and proteomics would be required. This has already been widely studied in other *F. oxysporum* f. spp. and provides evidence that genes and subsequent proteins are directly related to pathogenicity or whether they have been incorrectly identified as putative effectors (Rep *et al.*, 2005; Houterman *et al.*, 2009; Gawehns *et al.*, 2014). This is important for establishing the molecular mechanisms for pathogenicity in different races and could lead to better understanding of the pathogen-host interactions taking place. Using effectors to understand plant – pathogen gene interactions could also lead to more robust plant resistance to FOP, which is less likely to breakdown with the evolution of new races.

Additional future work should include whole genome sequencing of a wider selection of isolates from FOP races to confirm the presence of FOP specific and race specific putative effectors. In addition, pathogenic root rot isolates should also be whole genome sequenced and compared to FOP races to determine similarity and differences, and to understand the different mechanisms of infection between them. Preliminary analysis (Dr

A. Taylor, Warwick Crop Centre) suggested that there were shared mimp related genes in root rot and FOP isolates, which appear absent from other f. spp., suggesting possible host adaptation. Comparing genome sequences of root rot *F. oxysporum* isolates to the published *F. solani* whole genome could make important strides in the determination of pathogenicity in Fusarium root rot. As host resistance mechanisms for root rot in pea are not fully understood, this could provide insights into developing pea cultivars with broad resistance to multiple root rot *Fusarium* species found in the root rot complex.

The identity of FOP isolates representing the three different races as classified by *TEF* phylogeny and *SIX* gene profile was determined using differential pea cultivars. The identity of race 1 isolates was confirmed, but the distinction between races 2 and 5 was inconclusive, and suggested that the race 5 isolates might be a variation of race 2. However, from the differential cultivars used this was not possible to determine. Putative race 2 and 5 had a similar *SIX* gene profile (except for the absence of *SIX9* and *SIX14* in race 5), and it was hypothesised that race 5 isolates were in a different lineage to race 1 and 2 (Chapter 2), suggesting that FOP races are polyphyletic and that race 5 emerged as a result of convergent evolution. This could have been the result of race 5 acquiring only some of the *SIX* genes from race 2 via horizontal gene transfer. Polyphyletic origins of races have been used to provide evidence of horizontal gene transfer in other f. spp., for example, FOL and f. sp. *cubense* (Czislowski *et al.*, 2018; van Dam & Rep, 2017). However, the polyphyletic nature of FOP races cannot be determined from this study as it is unclear whether the isolates are race 2 or race 5. In the future, additional pea differential cultivars and isolates would need to be used to distinguish between these two putative races. Whole genome sequencing of further isolates in these putative races would confirm the similarities or differences between them and support the conclusions of the traditional race-typing pathogenicity tests. Comparing the genomes of the putative race 5 isolate in this study to a previously whole genome sequenced FOP race 5 isolate (Williams *et al.*, 2016), and including it in a differential pathogenicity test to confirm race type, would aid the identification of these isolates.

Although this work enhances our knowledge of *Fusarium* diseases in pea, especially Fusarium wilt, there is definitely scope to further our understanding with additional research in the future. As mentioned previously, completing a substantial survey of *Fusarium* diseases in pea in the UK, in addition to other pathogens, with increased

sampling sites, multiple sample times in the growing season, multiple growing seasons and consistency in sample preparation from the plant, would help to improve grower advice as to which fields to choose (those with a low disease index) and improve identification of wilt and root rot symptoms in the field to aid decisions about variety selection. In addition, this work would benefit from additional whole genome sequencing of FOP races from different clades in the *TEF* phylogeny, and isolates causing root rot, to determine the differences between pathogenicity factors in these pathogen types. Knock out studies of putative effector genes identified in these genomes would determine the impact they have on pathogenicity. Finally, additional FOP isolates should be race typed using a greater selection of the pea differential cultivars to confirm discrepancies in race type.

In summary, this work represents a considerable advance in the understanding of *Fusarium* diseases in pea. It is the first study of the distribution and pathogenicity of *F. oxysporum* isolates in the UK, and the first genomic study of multiple FOP races. Due to the risk of resistance breakdown in pea by the emergence of new FOP races, it is important to understand the molecular mechanisms of pathogenicity, in order to develop more robust resistant cultivars. This work could lead to the development of molecular diagnostics using the presence/absence of FOP specific and race specific genes, to help growers rapidly identify disease and plan crop rotations and cultivar choice for future growing seasons.

References

Adedamola, A. O. (2014) Identifying and characterising Fusarium species in pea. MSc Thesis, School of Life Sciences, University of Warwick

Alabouvette, C., Olivain, C., Migheli, Q. and Steinberg, C. (2009) Microbiological control of soil-borne phytopathogenic fungi with special emphasis on wilt-inducing Fusarium oxysporum. *New Phytologist,* **184,** 529-544.

Amatulli, M. T., Spadaro, D., Gullino, M. L. and Garibaldi, A. (2012) Molecular identification of Fusarium spp. associated with bakanae disease of rice in Italy and assessment of their pathogenicity. *Plant Pathology,* **61,** 820-820.

Anders, S. and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biology,* **11,** R106.

Armitage, A. D., Taylor, A., Sobczyk, M. K., Baxter, L., Greenfield, B. P. J., Bates, H. J., Wilson, F., Jackson, A. C., Ott, S., Harrison, R. J. and Clarkson, J. P. (2018) Characterisation of pathogen-specific regions and novel effector candidates in Fusarium oxysporum f. sp. cepae. *Scientific Reports,* **8,** 13530.

Armstrong, G. M. and Armstrong, J. K. (1974) Races of Fusarium oxysporum f.sp pisi, causal agents of wilt of pea. *Phytopathology,* **64,** 849-857.

Aronesty, E. (2013) Comparison of sequencing utility programs. *The Open Bioinformatics Journal,* **7,** 1-8.

Avalos, J. and Estrada, A. F. (2010) Regulation by light in Fusarium. *Fungal Genetics and Biology,* **47,** 930-938.

Baayen, R. P., O'Donnell, K., Bonants, P. J. M., Cigelnik, E., Kroon, L., Roebroeck, E. J. A. and Waalwijk, C. (2000) Gene genealogies and AFLP analyses in the Fusarium oxysporum complex identify monophyletic and nonmonophyletic formae speciales causing wilt and rot disease. *Phytopathology,* **90,** 891-900.

Bairoch, A. and Apweiler, R. (2000) The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Research,* **28,** 45-48.

Balmas, V., Migheli, Q., Scherm, B., Garau, P., O'Donnell, K., Ceccherelli, G., Rang, S. and Geiser, D. M. (2010) Multilocus phylogenetics show high levels of endemic fusaria inhabiting Sardinian soils (Tyrrhenian Islands). *Mycologia,* **102,** 803-812.

Bani, M., Rispail, N., Evidente, A., Rubiales, D. and Cimmino, A. (2014) Identification of the main toxins isolated from Fusarium oxysporum f. sp. pisi race 2 and their relation with isolates' pathogenicity. *Journal of agricultural and food chemistry,* **62,** 2574-2580.

Bani, M., Rubiales, D. and Rispail, N. (2012) A detailed evaluation method to identify sources of quantitative resistance to Fusarium oxysporum f. sp pisi race 2 within a Pisum spp. germplasm collection. *Plant Pathology,* **61,** 532-542.

Barilli, E., Cobos, M. J. and Rubiales, D. (2016) Clarification on host range of Didymella pinodes the causal agent of pea Ascochyta blight. *Frontiers in Plant Science,* **7,** 592.

Bayraktar, H. (2010) Genetic diversity and population structure of Fusarium oxysporum f. sp cepae, the causal agent of Fusarium basal plate rot on onion, using RAPD markers. *Tarim Bilimleri Dergisi-Journal of Agricultural Sciences,* **16,** 139-149.

Bergemann, M., Lespinet, O., Ben M'Barek, S., Daboussi, M. J. and Dufresne, M. (2008) Genome-wide analysis of the Fusarium oxysporum mimp family of MITES and mobilization of both native and *de novo* created mimps. *Journal of Molecular Evolution,* **67,** 631-642.

Bhattacharjee, R. B., Singh, A. and Mukhopadhyay, S. N. (2008) Use of nitrogen-fixing bacteria as biofertiliser for non-legumes: prospects and challenges. *Applied Microbiology and Biotechnology,* **80,** 199-209.

Biddle, A. J. and Cattlin, N. D. (2007) Pests, Diseases, and Disorders of Peas and Beans: A Colour Handbook. Academic Press.

Bodah, E. T., Porter, L. D., Chaves, B. and Dhingra, A. (2016) Evaluation of pea accessions and commercial cultivars for Fusarium root rot resistance. *Euphytica,* **208,** 63-72.

Boratyn, G. M., Camacho, C., Cooper, P. S., Coulouris, G., Fong, A., Ma, N., Madden, T. L., Matten, W. T., McGinnis, S. D., Merezhuk, Y., Raytselis, Y., Sayers, E. W., Tao, T., Ye, J. and Zaretskaya, I. (2013) BLAST: a more efficient report with usability improvements. *Nucleic Acids Research,* **41,** W29-W33.

Boye, J., Zare, F. and Pletch, A. (2010) Pulse proteins: Processing, characterization, functional properties and applications in food and feed. *Food Research International,* **43,** 414-431.

Buxton, E. W. (1955) Fusarium diseases of peas. *Transactions of the British Mycological Society,* **38,** 309.

Cao, L. X., Blekemolen, M. C., Tintor, N., Cornelissen, B. J. C. and Takken, F. L. W. (2018) The Fusarium oxysporum Avr2-Six5 effector pair alters plasmodesmatal exclusion selectivity to facilitate cell-to-cell movement of Avr2. *Molecular Plant,* **11,** 691-705.

Carlile, M. J. (1956) A study of the factors influencing non-genetic variation in a strain of Fusarium oxysporum. *Journal of General Microbiology,* **14,** 643-654.

Castillejo, M. A., Bani, M. and Rubiales, D. (2015) Understanding pea resistance mechanisms in response to Fusarium oxysporum through proteomic analysis. *Phytochemistry,* **115,** 44-58.

Chakrabarti, A., Rep, M., Wang, B., Ashton, A., Dodds, P. and Ellis, J. (2011) Variation in potential effector genes distinguishing Australian and non-Australian isolates of the cotton wilt pathogen Fusarium oxysporum f.sp vasinfectum. *Plant Pathology,* **60,** 232-243.

Chen, H. and Boutros, P. C. (2011) VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *Bmc Bioinformatics,* **12,** 35.

Chittem, K., Mathew, F. M., Gregoire, M., Lamppa, R. S., Chang, Y., Markell, S. G., Bradley, C. A., Barasubiye, T. and Goswami, R. S. (2015) Identification and characterization of Fusarium spp. associated with root rots of field pea in North Dakota. *European Journal of Plant Pathology,* **143,** 641-649.

Clarkson, J. D. S. (1978) Pathogenicity of Fusarium spp. associated with foot-rots of peas and beans. *Plant Pathology,* **27,** 110-117.

Coleman, J. J., Rounsley, S. D., Rodriguez-Carres, M., Kuo, A., Wasmann, C. C., Grimwood, J., Schmutz, J., Taga, M., White, G. J., Zhou, S., Schwartz, D. C., Freitag, M., Ma, L.-j., Danchin, E. G. J., Henrissat, B., Coutinho, P. M., Nelson, D. R., Straney, D., Napoli, C. A., Barker, B. M., Gribskov, M., Rep, M., Kroken, S., Molnar, I., Rensing, C., Kennell, J. C., Zamora, J., Farman, M. L., Selker, E. U., Salamov, A., Shapiro, H., Pangilinan, J., Lindquist, E., Lamers, C., Grigoriev, I. V., Geiser, D. M., Covert, S. F., Temporini, E. and VanEtten, H. D. (2009) The genome of Nectria haematococca: Contribution of supernumerary chromosomes to gene expansion. *Plos Genetics,* **5,** e1000618.

Coleman, J. J., Wasmann, C. C., Usami, T., White, G. J., Temporini, E. D., McCluskey, K. and VanEtten, H. D. (2011) Characterization of the gene encoding pisatin

demethylase (FoPDA1) in Fusarium oxysporum. *Molecular Plant-Microbe Interactions,* **24,** 1482-1491.

Covey, P. A., Kuwitzky, B., Hanson, M. and Webb, K. M. (2014) Multilocus analysis using putative fungal effectors to describe a population of Fusarium oxysporum from sugar beet. *Phytopathology,* **104,** 886-896.

Czislowski, E., Fraser-Smith, S., Zander, M., O'Neill, W. T., Meldrum, R. A., Tran-Nguyen, L. T. T., Batley, J. and Aitken, E. A. B. (2018) Investigation of the diversity of effector genes in the banana pathogen, Fusarium oxysporum f. sp cubense, reveals evidence of horizontal gene transfer. *Molecular Plant Pathology,* **19,** 1155-1171.

Dahl, W. J., Foster, L. M. and Tyler, R. T. (2012) Review of the health benefits of peas (Pisum sativum L.). *British Journal of Nutrition,* **108,** S3-S10.

de Sain, M. and Rep, M. (2015) The role of pathogen-secreted proteins in fungal vascular wilt diseases. *International Journal of Molecular Sciences,* **16,** 23970-23993.

Dean, R., Van Kan, J. A. L., Pretorius, Z. A., Hammond-Kosack, K. E., Di Pietro, A., Spanu, P. D., Rudd, J. J., Dickman, M., Kahmann, R., Ellis, J. and Foster, G. D. (2012) The top 10 fungal pathogens in molecular plant pathology. *Molecular Plant Pathology,* **13,** 414-430.

Di Pietro, A., Garcia-Maceira, F. I., Meglecz, E. and Roncero, M. I. G. (2001) A MAP kinase of the vascular wilt fungus Fusarium oxysporum is essential for root penetration and pathogenesis. *Molecular Microbiology,* **39,** 1140-1152.

Di Pietro, A., Madrid, M. P., Caracuel, Z., Delgado-Jarana, J. and Roncero, M. I. G. (2003) Fusarium oxysporum: exploring the molecular arsenal of a vascular wilt fungus. *Molecular Plant Pathology,* **4,** 315-325.

Dissanayake, M., Kashima, R., Tanaka, S. and Ito, S. (2009) Pathogenic variation and molecular characterization of Fusarium species isolated from wilted Welsh onion in Japan. *Journal of General Plant Pathology,* **75,** 37-45.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics,* **29,** 15-21.

Dufresne, M., Hua-Vant, A., el Wahab, H. A., Ben M'Barek, S., Vasnier, C., Teysset, L., Kema, G. H. J. and Daboussi, M. J. (2007) Transposition of a fungal miniature inverted-repeat transposable element through the action of a TH-like transposase. *Genetics,* **175,** 441-452.

Dyer, P. S. and Ingram, D. S. (1990) An improved test for evaluating the pathogencity of isolates of Fusarium solani f.sp pisi on pea. *Annals of Applied Biology,* **117,** 469-472.

Edel, V., Steinberg, C., Gautheron, N. and Alabouvette, C. (1997) Populations of nonpathogenic Fusarium oxysporum associated with roots of four plant species compared to soilborne populations. *Phytopathology,* **87,** 693-697.

Eurofins (2016) Eurofins Oligo Analysis Tool. http://www.operon.com/tools/oligo-analysis-tool.aspx (Accessed 2016).

FAOSTAT (2016) FAOSTAT Crops: Peas, green. http://www.fao.org/faostat/en/#data/QC (Accessed 2018).

Felsenstein, J. (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution; International Journal of Organic Evolution,* **39,** 783-791.

Feng, J., Hwang, R., Chang, K. F., Hwang, S. F., Strelkov, S. E., Gossen, B. D., Conner, R. L. and Turnbull, G. D. (2010) Genetic variation in Fusarium avenaceum causing root rot on field pea. *Plant Pathology,* **59,** 845-852.

Fernandez, M. R. (2007) Fusarium populations in roots of oilseed and pulse crops grown in eastern Saskatchewan. *Canadian Journal of Plant Science,* **87,** 945-952.

Finn, R. D., Clements, J. and Eddy, S. R. (2011) HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research,* **39,** W29-W37.

Fraser-Smith, S., Czislowski, E., Meldrum, R. A., Zander, M., O'Neill, W., Balali, G. R. and Aitken, E. A. B. (2014) Sequence variation in the putative effector gene SIX8 facilitates molecular differentiation of Fusarium oxysporum f. sp cubense. *Plant Pathology,* **63,** 1044-1052.

Fuchs, J. G., MoenneLoccoz, Y. and Defago, G. (1997) Nonpathogenic Fusarium oxysporum strain Fo47 induces resistance to Fusarium wilt in tomato. *Plant Disease,* **81,** 492-496.

Gadagkar, S. R., Rosenberg, M. S. and Kumar, S. (2005) Inferring species phylogenies from multiple genes: Concatenated sequence tree versus consensus gene tree. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution,* **304B,** 64-74.

Garibaldi, A., Gilardi, G. and Gullino, M. L. (2004) Seed transmission of Fusarium oxysporum f.sp. lactucae. *Phytoparasitica,* **32,** 61-65.

Gawehns, F., Houterman, P. M., Ichou, F. A., Michielse, C. B., Hijdra, M., Cornelissen, B. J., Rep, M. and Takken, F. L. (2014) The Fusarium oxysporum effector Six6 contributes to virulence and suppresses I-2-mediated cell death. *Molecular Plant-Microbe Interactions,* **27,** 336-348.

Geiser, D. M., Jimenez-Gasco, M. D., Kang, S. C., Makalowska, I., Veeraraghavan, N., Ward, T. J., Zhang, N., Kuldau, G. A. and O'Donnell, K. (2004) FUSARIUM-ID v. 1.0: A DNA sequence database for identifying Fusarium. *European Journal of Plant Pathology,* **110,** 473-479.

Gibriel, H. A. Y., Thomma, B. and Seidl, M. F. (2016) The age of effectors: genome-based discovery and applications. *Phytopathology,* **106,** 1206-1212.

Giordano, F., Aigrain, L., Quail, M. A., Coupland, P., Bonfield, J. K., Davies, R. M., Tischler, G., Jackson, D. K., Keane, T. M., Li, J., Yue, J. X., Liti, G., Durbin, R. and Ning, Z. M. (2017) De novo yeast genome assemblies from MinION, PacBio and MiSeq platforms. *Scientific Reports,* **7,** 3935.

Gomes, E. V., Costa, M. D., de Paula, R. G., de Azevedo, R. R., da Silva, F. L., Noronha, E. F., Ulhoa, C. J., Monteiro, V. N., Cardoza, R. E., Gutierrez, S. and Silva, R. N. (2015) The Cerato-Platanin protein Epl-1 from Trichoderma harzianum is involved in mycoparasitism, plant resistance induction and self cell wall protection. *Scientific Reports,* **5,** 17998.

Gordon, T. R. (2017) Fusarium oxysporum and the Fusarium Wilt Syndrome. *Annual Review of Phytopathology, Vol 55,* **55,** 23-39.

Grajal-martin, M. J., Simon, C. J. and Muehlbauer, F. J. (1993) Use of random amplified polymorphic DNA (RAPD) to characterize race 2 of Fusarium oxysporum f. sp. pisi. *Phytopathology,* **83,** 612-614.

Green, K. and Foster, C. (2005) Give peas a chance: Transformations in food consumption and production systems. *Technological Forecasting and Social Change,* **72,** 663-679.

Grunwald, N. J., Coffman, V. A. and Kraft, J. M. (2003) Sources of partial resistance to Fusarium root rot in the Pisum core collection. *Plant Disease,* **87,** 1197-1200.

Guo, L. J., Han, L. J., Yang, L. Y., Zeng, H. C., Fan, D. D., Zhu, Y. B., Feng, Y., Wang, G. F., Peng, C. F., Jiang, X. T., Zhou, D. J., Ni, P. X., Liang, C. C., Liu, L., Wang, J., Mao, C., Fang, X. D., Peng, M. and Huang, J. S. (2014) Genome and

transcriptome analysis of the fungal pathogen Fusarium oxysporum f. sp cubense causing banana vascular wilt disease. *Plos One,* **9,** e95543.
Gurevich, A., Saveliev, V., Vyahhi, N. and Tesler, G. (2013) QUAST: quality assessment tool for genome assemblies. *Bioinformatics,* **29,** 1072-1075.
Haas, B. (2007) TransposonPSI. http://transposonpsi.sourceforge.net (Accessed 2018).
Haglund, W. A. (1989) A rapid method for inoculating pea-seedlings with fusarium-oxysporum f-sp pisi. *Plant Disease,* **73,** 457-458.
Haglund, W. A. and Kraft, J. M. (1970) Fusarium oxysporum f. pisi race 5. *Phytopathology,* **60,** 1861-1862.
Haglund, W. A. and Kraft, J. M. (1979) Fusarium-oxysporum f-sp pisi, race-6 - occurrence and distribution. *Phytopathology,* **69,** 818-820.
Han, Y. X., Gao, S. G., Muegge, K., Zhang, W. and Zhou, B. (2015) Advanced applications of RNA sequencing and challenges. *Bioinformatics and Biology Insights,* **9,** 29-46.
Hoff, K. J., Lange, S., Lomsadze, A., Borodovsky, M. and Stanke, M. (2016) BRAKER1: Unsupervised RNA-seq-based genome annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics,* **32,** 767-769.
Houterman, P. M., Cornelissen, B. J. C. and Rep, M. (2008) Suppression of plant resistance gene-based immunity by a fungal effector. *PLoS Pathogens,* **4,** e1000061.
Houterman, P. M., Ma, L., van Ooijen, G., de Vroomen, M. J., Cornelissen, B. J. C., Takken, F. L. W. and Rep, M. (2009) The effector protein Avr2 of the xylem-colonizing fungus Fusarium oxysporum activates the tomato resistance protein I-2 intracellularly. *Plant Journal,* **58,** 970-978.
Houterman, P. M., Speijer, D., Dekker, H. L., de Koster, C. G., Cornelissen, B. J. C. and Rep, M. (2007) The mixed xylem sap proteome of Fusarium oxysporum-infected tomato plants. *Molecular Plant Pathology,* **8,** 215-221.
Hua-Van, A., Langin, T. and Daboussi, M. J. (2001) Evolutionary history of the impala transposon in Fusarium oxysporum. *Molecular Biology and Evolution,* **18,** 1959-1969.
Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., Bravo, H. C., Davis, S., Gatto, L., Girke, T., Gottardo, R., Hahne, F., Hansen, K. D., Irizarry, R. A., Lawrence, M., Love, M. I., MacDonald, J., Obenchain, V., Oles, A. K., Pages, H., Reyes, A., Shannon, P., Smyth, G. K., Tenenbaum, D., Waldron, L. and Morgan, M. (2015) Orchestrating high-throughput genomic analysis with Bioconductor. *Nature Methods,* **12,** 115-121.
Hwang, S. F. and Chang, K. F. (1989) Incidence and severity of root rot disease complex of field pea in northeastern Alberta in 1988. *Canadian Plant Disease Survey,* **2,** 139-141.
Infantino, A., Kharrat, M., Riccioni, L., Coyne, C. J., McPhee, K. E. and Grunwald, N. J. (2006) Screening techniques and sources of resistance to root diseases in cool season food legumes. *Euphytica,* **147,** 201-221.
Ivic, D. (2014) Pathogenicity and potential toxigenicity of seed-borne Fusarium species on soybean and pea. *Journal of Plant Pathology,* **96,** 541-551.
Jones, F. R. and Linford, M. B. (1925) Pea disease survey in Wisconsin. *Wisconsin Agricultural Experiment Station Research Bulletin,* **64,** 31.
Jones, J. D. G. and Dangl, J. L. (2006) The plant immune system. *Nature,* **444,** 323-329.
Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W. Z., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas,

A., Scheremetjew, M., Yong, S. Y., Lopez, R. and Hunter, S. (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics,* **30,** 1236-1240.

Jonkers, W., Rodrigues, C. D. A. and Rep, M. (2009) Impaired colonization and infection of tomato roots by the Delta frp1 mutant of Fusarium oxysporum correlates with reduced CWDE gene expression. *Molecular Plant-Microbe Interactions,* **22,** 507-518.

Kamoun, S. (2007) Groovy times: filamentous pathogen effectors revealed. *Current Opinion in Plant Biology,* **10,** 358-365.

Kamoun, S. (2009) The secretome of plant-associated fungi and oomycetes. In: *Deising H.B. (eds) Plant Relationships. The Mycota (A Comprehensive Treatise on Fungi as Experimental Systems for Basic and Applied Research).* Berlin, Heidelberg: Springer, vol 5.

Khan, F. and Khan, M. I. (2011) Fungal lectins: current molecular and biochemical perspectives. *International Journal of Biological Chemistry,* **5,** 1-20.

Kimura, M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution,* **16,** 111-120.

Kraft, J. and Haglund, W. (1978) A reappraisal of the race classification of Fusarium oxysporum f. sp. pisi. *Phytopathology,* **68,** 273-276.

Kraft, J. M. (1994) Fusarium-wilt of peas (a review). *Agronomie,* **14,** 561-567.

Kraft, J. M., Haware, M. P., Jiménez-Díaz, R. M., Bayaa, B. and Harrabi, M. (1994) Screening techniques and sources of resistance to root rots and wilts in cool season food legumes. *Euphytica,* **73,** 27-39.

Kraft, J. M. and Kaiser, W. J. (1993) Screening for disease resistance in pea. In: Breeding for Stress Tolerance in Cool-Season Food Legumes. Chichester, UK: John Wiley & Sons.

Kraft, J. M., Muehlbauer, F. J., Cook, R. J. and Entemann, F. M. (1974) The reappearance of common wilt of peas in eastern Washington. *Plant Disease Reports,* **58,** 62-64.

Kraft, J. M. and Pfleger, F. L. (2001) Compendium of Pea Diseases and Pests (2nd ed.). St. Paul, MN. : The American Phytopathological Society.

Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E. L. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology,* **305,** 567-580.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S. J. and Marra, M. A. (2009) Circos: An information aesthetic for comparative genomics. *Genome Research,* **19,** 1639-1645.

Kumar, S., Stecher, G. and Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. *Molecular Biology and Evolution,* **33,** 1870-1874.

Leslie, J. F. (1993) Fungal vegetative compatibility. *Annual Review of Phytopathology,* **31,** 127-150.

Leslie, J. F. and Summerell, B. A. (2006) The *Fusarium* Laboratory Manual. Blackwell Publishing Ltd, Ames, IA.

Leslie, J. F., Zeller, K. A. and Summerell, B. A. (2001) Icebergs and species in populations of Fusarium. *Physiological and Molecular Plant Pathology,* **59,** 107-117.

Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M. and Henrissat, B. (2013) Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnology for Biofuels,* **6,** 41.

Li, L., Stoeckert, C. J. and Roos, D. S. (2003) OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Research,* **13,** 2178-2189.

Li, X. M., Liu, C. J., Chakraborty, S., Manners, J. M. and Kazan, K. (2008) A simple method for the assessment of crown rot disease severity in wheat seedlings inoculated with Fusarium pseudograminearum. *Journal of Phytopathology,* **156,** 751-754.

Lievens, B., Houterman, P. M. and Rep, M. (2009) Effector gene screening allows unambiguous identification of Fusarium oxysporum f. sp. lycopersici races and discrimination from other formae speciales. *FEMS Microbiol Lett,* **300,** 201-215.

Linfield, C. A. (1994) An investigation into the detection, identification, differentiation and pathogenicity of Fusarium oxysporum f.sp. narcissi. PhD Thesis, Warwick HRI, University of Reading.

Loman, N. J., Quick, J. and Simpson, J. T. (2015) A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nature Methods,* **12,** 733-735.

Lombard, V., Ramulu, H. G., Drula, E., Coutinho, P. M. and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research,* **42,** D490-D495.

Lysøe, E., Bone, K. R. and Klemsdal, S. S. (2008) Identification of up-regulated genes during zearalenone biosynthesis in Fusarium. *European Journal of Plant Pathology,* **122,** 505-516.

Ma, L. J. (2014) Horizontal chromosome transfer and rational strategies to manage Fusarium vascular wilt diseases. *Molecular Plant Pathology,* **15,** 962-962.

Ma, L. J., Geiser, D. M., Proctor, R. H., Rooney, A. P., O'Donnell, K., Trail, F., Gardiner, D. M., Manners, J. M. and Kazan, K. (2013) Fusarium pathogenomics. *Annual Review of Microbiology,* **67,** 399-416.

Ma, L. J., van der Does, H. C., Borkovich, K. A., Coleman, J. J., Daboussi, M. J., Di Pietro, A., Dufresne, M., Freitag, M., Grabherr, M., Henrissat, B., Houterman, P. M., Kang, S., Shim, W. B., Woloshuk, C., Xie, X. H., Xu, J. R., Antoniw, J., Baker, S. E., Bluhm, B. H., Breakspear, A., Brown, D. W., Butchko, R. A. E., Chapman, S., Coulson, R., Coutinho, P. M., Danchin, E. G. J., Diener, A., Gale, L. R., Gardiner, D. M., Goff, S., Hammond-Kosack, K. E., Hilburn, K., Hua-Van, A., Jonkers, W., Kazan, K., Kodira, C. D., Koehrsen, M., Kumar, L., Lee, Y. H., Li, L. D., Manners, J. M., Miranda-Saavedra, D., Mukherjee, M., Park, G., Park, J., Park, S. Y., Proctor, R. H., Regev, A., Ruiz-Roldan, M. C., Sain, D., Sakthikumar, S., Sykes, S., Schwartz, D. C., Turgeon, B. G., Wapinski, I., Yoder, O., Young, S., Zeng, Q. D., Zhou, S. G., Galagan, J., Cuomo, C. A., Kistler, H. C. and Rep, M. (2010) Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. *Nature,* **464,** 367-373.

Ma, L. S., Houterman, P. M., Gawehns, F., Cao, L. X., Sillo, F., Richter, H., Clavijo-Ortiz, M. J., Schmidt, S. M., Boeren, S., Vervoort, J., Cornelissen, B. J. C., Rep, M. and Takken, F. L. W. (2015) The AVR2-SIX5 gene pair is required to activate I-2-mediated immunity in tomato. *New Phytologist,* **208,** 507-518.

Maguire, K. (Year) The underground world of peas; the interactions between roots and microbes, Warwick Crop Centre Seminar Series, Warwick Crop Centre, The University of Warwick.

Marinangeli, C. P. F., Kassis, A. N. and Jones, P. J. H. (2009) Glycemic responses and sensory characteristics of whole yellow pea flour added to novel functional foods. *Journal of Food Science,* **74,** S385-S389.

McPhee, K. E., Tullu, A., Kraft, J. M. and Muehlbauer, F. J. (1999) Resistance to Fusarium wilt race 2 in the Pisum core collection. *Journal of the American Society for Horticultural Science,* **124,** 28-31.

Meldrum, R. A., Fraser-Smith, S., Tran-Nguyen, L. T. T., Daly, A. M. and Aitken, E. A. B. (2012) Presence of putative pathogenicity genes in isolates of Fusarium oxysporum f. sp. cubense from Australia. *Australasian Plant Pathology,* **41,** 551-557.

Merzoug, A., Belabid, L., Youcef-Benkada, M., Benfreha, F. and Bayaa, B. (2014) Pea Fusarium wilt races in western Algeria. *Plant Protection Science,* **50,** 70-77.

Michielse, C. B., van Wijk, R., Reijnen, L., Manders, E. M. M., Boas, S., Olivain, C., Alabouvette, C. and Rep, M. (2009) The nuclear protein Sge1 of Fusarium oxysporum is required for parasitic growth. *PLoS Pathogens,* **5,** e1000637.

Moore, K. L., Rodriguez-Ramiro, I., Jones, E. R., Jones, E. J., Rodriguez-Celma, J., Halsey, K., Domoney, C., Shewry, P. R., Fairweather-Tait, S. and Balk, J. (2018) The stage of seed development influences iron bioavailability in pea (Pisum sativum L.). *Scientific Reports,* **8,** 6865.

Moretti, A. (2009) Taxonomy of Fusarium genus: A continuous fight between lumpers and splitters. *Proceedings for Natural Sciences, Matica Srpska***,** 7-13.

NCBI (2018) Taxonomy browser: Fusarium. https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?mode=Info&id=5506&lvl=3&lin=f&keep=1&srchmode=1&unlock (Accessed 2018).

Nelson, P. E., Dignani, M. C. and Anaissie, E. J. (1994) Taxonomy, biology, and clinical aspects of fusarium species. *Clinical Microbiology Reviews,* **7,** 479-504.

Neumann, S. and Xue, A. G. (2003) Reactions of field pea cultivars to four races of Fusarium oxysporum f. sp pisi. *Canadian Journal of Plant Science,* **83,** 377-379.

Newman, M. A., Sundelin, T., Nielsen, J. T. and Erbs, G. (2013) MAMP (microbe-associated molecular pattern) triggered immunity in plants. *Frontiers in Plant Science,* **4,** 139.

Nicholson, P., Chandler, E., Draeger, R. C., Gosman, N. E., Simpson, D. R., Thomsett, M. and Wilson, A. H. (2003) Molecular tools to study epidemiology and toxicology of fusarium head blight of cereals. *European Journal of Plant Pathology,* **109,** 691-703.

Nino-Sanchez, J., Casado-Del Castillo, V., Tello, V., De Vega-Bartol, J. J., Ramos, B., Sukno, S. A. and Minguez, J. M. D. (2016) The FTF gene family regulates virulence and expression of SIX effectors in Fusarium oxysporum. *Molecular Plant Pathology,* **17,** 1124-1139.

Nirmaladevi, D., Venkataramana, M., Srivastava, R. K., Uppalapati, S. R., Gupta, V. K., Yli-Mattila, T., Tsui, K. M. C., Srinivas, C., Niranjana, S. R. and Chandra, N. S. (2016) Molecular phylogeny, pathogenicity and toxigenicity of Fusarium oxysporum f. sp lycopersici. *Scientific Reports,* **6,** 21367.

Niu, X. W., Zhao, X. Q., Ling, K. S., Levi, A., Sun, Y. Y. and Fan, M. (2016) The FonSIX6 gene acts as an avirulence effector in the Fusarium oxysporum f. sp niveum - watermelon pathosystem. *Scientific Reports,* **6,** 28146.

Nodvig, C. S., Nielsen, J. B., Kogle, M. E. and Mortensen, U. H. (2015) A CRISPR-Cas9 system for genetic engineering of filamentous fungi. *Plos One,* **10,** e0133085.

Nurk, S., Bankevich, A., Antipov, D., Gurevich, A. A., Korobeynikov, A., Lapidus, A., Prjibelski, A. D., Pyshkin, A., Sirotkin, A., Sirotkin, Y., Stepanauskas, R., Clingenpeel, S. R., Woyke, T., McLean, J. S., Lasken, R., Tesler, G., Alekseyev, M. A. and Pevzner, P. A. (2013) Assembling single-cell genomes and mini-

metagenomes from chimeric MDA products. *Journal of Computational Biology,* **20,** 714-737.

O'Donnell, K., Kistler, H. C., Cigelnik, E. and Ploetz, R. C. (1998) Multiple evolutionary origins of the fungus causing Panama disease of banana: Concordant evidence from nuclear and mitochondrial gene genealogies. *Proceedings of the National Academy of Sciences of the United States of America,* **95,** 2044-2049.

O'Donnell, K., Sarver, B. A. J., Brandt, M., Chang, D. C., Noble-Wang, J., Park, B. J., Sutton, D. A., Benjamin, L., Lindsley, M., Padhye, A., Geiser, D. M. and Ward, T. J. (2007) Phylogenetic diversity and microsphere array-based genotyping of human pathogenic fusaria, including isolates from the multistate contact lens - Associated US Keratitis outbreaks of 2005 and 2006. *Journal of Clinical Microbiology,* **45,** 2235-2248.

O'Donnell, K., Sutton, D. A., Rinaldi, M. G., Sarver, B. A. J., Balajee, S. A., Schroers, H. J., Summerbell, R. C., Robert, V., Crous, P. W., Zhang, N., Aoki, T., Jung, K., Park, J., Lee, Y. H., Kang, S., Park, B. and Geiser, D. M. (2010) Internet-accessible DNA sequence database for identifying Fusaria from human and animal Infections. *Journal of Clinical Microbiology,* **48,** 3708-3718.

O'Donnell, K. and Cigelnik, E. (1997) Two divergent intragenomic rDNA ITS2 types within a monophyletic lineage of the fungus Fusarium are nonorthologous. *Molecular Phylogenetics and Evolution,* **7,** 103-116.

Ochiai, A., Yamasaki, M., Mikami, B., Hashimoto, W. and Murata, K. (2010) Crystal structure of exotype alginate lyase Atu3025 from Agrobacterium tumefaciens. *Journal of Biological Chemistry,* **285,** 24519-24528.

Olivieri, F., Zanetti, M. E., Oliva, C. R., Covarrubias, A. A. and Casalongue, C. A. (2002) Characterization of an extracellular serine protease of Fusarium eumartii and its action on pathogenesis related proteins. *European Journal of Plant Pathology,* **108,** 63-72.

Ondrej, M., Dostalova, R. and Trojam, R. (2008) Evaluation of virulence of Fusarium solani isolates on pea. *Plant Protection Science,* **44,** 9-18.

Ospina-Giraldo, M. D., Griffith, J. G., Laird, E. W. and Mingora, C. (2010) The CAZyome of Phytophthora spp.: A comprehensive analysis of the gene complement coding for carbohydrate-active enzymes in species of the genus Phytophthora. *BMC Genomics,* **11,** 525.

Oyarzun, P., Gerlagh, M. and Hoogland, A. E. (1993) Pathogenic fungi involved in root rot of peas in the Netherlands and their physiological specialization. *Netherlands Journal of Plant Pathology,* **99,** 23-33.

Ozer, N., Koycu, N. D., Chilosi, G. and Magro, P. (2004) Resistance to fusarium basal rot of onion in greenhouse and field and associated expression of antifungal compounds. *Phytoparasitica,* **32,** 388-394.

Pandey, A. and Mann, M. (2000) Proteomics to study genes and genomes. *Nature,* **405,** 837-846.

Parry, D. W., Jenkinson, P. and McLeod, L. (1995) Fusarium ear blight (scab) in small grain cereals - a review. *Plant Pathology,* **44,** 207-238.

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. and Kingsford, C. (2017) Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods,* **14,** 417-419.

Peoples, M. B., Brockwell, J., Herridge, D. F., Rochester, I. J., Alves, B. J. R., Urquiaga, S., Boddey, R. M., Dakora, F. D., Bhattarai, S., Maskey, S. L., Sampet, C., Rerkasem, B., Khan, D. F., Hauggaard-Nielsen, H. and Jensen, E. S. (2009) The

contributions of nitrogen-fixing crop legumes to the productivity of agricultural systems. *Symbiosis,* **48,** 1-17.

Perez-Nadales, E., Nogueira, M. F. A., Baldin, C., Castanheira, S., El Ghalid, M., Grund, E., Lengeler, K., Marchegiani, E., Mehrotra, P. V., Moretti, M., Naik, V., Oses-Ruiz, M., Oskarsson, T., Schaefer, K., Wasserstrom, L., Brakhage, A. A., Gow, N. A. R., Kahmann, R., Lebrun, M.-H., Perez-Martin, J., Di Pietro, A., Talbot, N. J., Toquin, V., Walther, A. and Wendland, J. (2014) Fungal model systems and the elucidation of pathogenicity determinants. *Fungal Genetics and Biology,* **70,** 42-67.

Persson, L., Bødker, L. and Larsson-Wikström, M. (1997) Prevalence and pathogenicity of foot and root rot pathogens of pea in southern Scandinavia. *Plant Disease,* **81,** 171-174.

Petersen, T. N., Brunak, S., von Heijne, G. and Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods,* **8,** 785-786.

PGRO (2015) PGRO vining pea growers guide. Processors and Growers Research Organisation

Ploetz, R. C. (2015) Management of Fusarium wilt of banana: A review with special reference to tropical race 4. *Crop Protection,* **73,** 7-15.

Porter, L. D. (2010) Identification of tolerance to Fusarium root rot in wild pea germplasm with high levels of partial resistance. *Pisum Genetics,* **42**.

Ramos, B., Alves-Santos, F. M., Garcia-Sdnehez, M. A., Martin-Rodrigues, N., Eslava, A. P. and Diaz-Minguez, J. M. (2007) The gene coding for a new transcription factor (ftfl) of Fusarium oxysporum is only expressed during infection of common bean. *Fungal Genetics and Biology,* **44,** 864-876.

Rep, M., Meijer, M., Houterman, P. M., van der Does, H. C. and Cornelissen, B. J. C. (2005) Fusarium oxysporum evades I-3-mediated resistance without altering the matching avirulence gene. *Molecular Plant-Microbe Interactions,* **18,** 15-23.

Rep, M., van der Does, H. C., Meijer, M., van Wijk, R., Houterman, P. M., Dekker, H. L., de Koster, C. G. and Cornelissen, B. J. C. (2004) A small, cysteine-rich protein secreted by Fusarium oxysporum during colonization of xylem vessels is required for I-3-mediated resistance in tomato. *Molecular Microbiology,* **53,** 1373-1383.

Rodriguezgalvez, E. and Mendgen, K. (1995) The infection process of Fusarium oxysporum in cotton root tips. *Protoplasma,* **189,** 61-72.

Ruan, J. (2016) Ultra-fast de novo assembler using long noisy reads. https://github.com/ruanjue/smartdenovo (Accessed 2018).

Sasaki, K., Nakahara, K., Tanaka, S., Shigyo, M. and Ito, S.-i. (2015) Genetic and pathogenic variability of Fusarium oxysporum f. sp cepae isolated from onion and Welsh onion in Japan. *Phytopathology,* **105,** 525-532.

Schmidt, S. M., Houterman, P. M., Schreiver, I., Ma, L., Amyotte, S., Chellappan, B., Boeren, S., Takken, F. L. and Rep, M. (2013) MITEs in the promoters of effector genes allow prediction of novel virulence genes in Fusarium oxysporum. *BMC Genomics,* **14,** 119.

Schmidt, S. M., Lukasiewicz, J., Farrer, R., van Dam, P., Bertoldo, C. and Rep, M. (2016) Comparative genomics of Fusarium oxysporum f. sp melonis reveals the secreted protein recognized by the Fom-2 resistance gene in melon. *New Phytologist,* **209,** 307-318.

Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., Chen, W. and Fungal Barcoding, C. (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings*

of the National Academy of Sciences of the United States of America, **109,** 6241-6246.

Sebastiani, M. S., Bagnaresi, P., Sestili, S., Biselli, C., Zechini, A., Orru, L., Cattivelli, L. and Ficcadenti, N. (2017) Transcriptome analysis of the melon-Fusarium oxysporum f.sp melonis race 1.2 pathosystem in susceptible and resistant plants. *Frontiers in Plant Science,* **8,** 362.

Selin, C., de Kievit, T. R., Belmonte, M. F. and Fernando, W. G. D. (2016) Elucidating the role of effectors in plant-fungal interactions: progress and challenges. *Frontiers in Microbiology,* **7,** 600.

Seminis (2013) Pea Assortment 2013. https://seminisuk.s3.amazonaws.com/wp-content/uploads/2015/08/Seminis_Pea_Pea-Assortment-Processing-Overview-2013-NO.pdf (Accessed: 2018).

Sharma, A., Rathour, R., Plaha, P., Katoch, V., Khalsa, G. S., Patial, V., Singh, Y. and Pathania, N. K. (2010) Induction of Fusarium wilt (Fusarium oxysporum f. sp pisi) resistance in garden pea using induced mutagenesis and in vitro selection techniques. *Euphytica,* **173,** 345-356.

Sharma, P. (2011) Alarming occurrence of Fusarium wilt disease in pea (Pisum sativum L.) cultivations of Jabalpur district in Central India revealed by an array of pathogenicity tests. *Agriculture and Biology Journal of North America,* **2,** 981-994.

Shelest, E. (2017) Transcription factors in fungi: TFome dynamics, three major families, and dual-specificity TFs. *Frontiers in Genetics,* **8,** 53.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. and Zdobnov, E. M. (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics,* **31,** 3210-3212.

Sista Kameshwar, A. K. and Qin, W. (2018) Comparative study of genome-wide plant biomass-degrading CAZymes in white rot, brown rot and soft rot fungi. *Mycology,* **9,** 93-105.

Skovgaard, K., Bødker, L. and Rosendahl, S. (2002) Population structure and pathogenicity of members of the Fusarium oxysporum complex isolated from soil and root necrosis of pea (Pisum sativum L.). *FEMS Microbiology Ecology,* **42,** 367-374.

Smit, A. F. A., Hubley, R. and Green, P. (2013-2015) RepeatMasker Open-4.0. http://www.repeatmasker.org (Accessed 2018).

Smith, L. J., Smith, M. K., Tree, D., O'Keefe, D. and Galea, V. J. (2008) Development of a small-plant bioassay to assess banana grown from tissue culture for consistent infection by Fusarium oxysporum f. sp cubense. *Australasian Plant Pathology,* **37,** 171-179.

Smith, S. A., Beaulieu, J. M. and Donoghue, M. J. (2010) An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. *Proceedings of the National Academy of Sciences,* **107,** 5897-5902.

Sorensen, L. Q., Lysoe, E., Larsen, J. E., Khorsand-Jamal, P., Nielsen, K. F. and Frandsen, R. J. N. (2014) Genetic transformation of Fusarium avenaceum by Agrobacterium tumefaciens mediated transformation and the development of a USER-Brick vector construction system. *Bmc Molecular Biology,* **15,** 15.

Sperschneider, J., Dodds, P. N., Gardiner, D. M., Manners, J. M., Singh, K. B. and Taylor, J. M. (2015) Advances and challenges in computational prediction of effectors from plant pathogenic fungi. *PLoS Pathogens,* **11,** e1004806.

Sperschneider, J., Gardiner, D. M., Dodds, P. N., Tini, F., Covarelli, L., Singh, K. B., Manners, J. M. and Taylor, J. M. (2016) EffectorP: predicting fungal effector proteins from secretomes using machine learning. *New Phytologist,* **210,** 743-761.

Stanke, M. and Morgenstern, B. (2005) AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Research,* **33,** W465-W467.

Stergiopoulos, I. and de Wit, P. J. G. M. (2009) Fungal effector proteins. *Annual Review of Phytopathology,* **47,** 233-263.

Summerell, B. A., Laurence, M. H., Liew, E. C. Y. and Leslie, J. F. (2010) Biogeography and phylogeography of Fusarium: a review. *Fungal Diversity,* **44,** 3-13.

Summerell, B. A. and Leslie, J. F. (2011) Fifty years of Fusarium: how could nine species have ever been enough? *Fungal Diversity,* **50,** 135-144.

Summerell, B. A., Salleh, B. and Leslie, J. F. (2003) A utilitarian approach to Fusarium identification. *Plant Disease,* **87,** 117-128.

Taheri, A., Hamel, C., Gan, Y. and Vujanovic, V. (2011) First report of Fusarium redolens from Saskatchewan and its comparative pathogenicity. *Canadian Journal of Plant Pathology,* **33,** 559-564.

Taheri, E. A., Chatterton, S., Foroud, N. A., Gossen, B. D. and McLaren, D. L. (2017) Identification and community dynamics of fungi associated with root, crown, and foot rot of field pea in western Canada. *European Journal of Plant Pathology,* **147,** 489-500.

Takken, F. and Rep, M. (2010) The arms race between tomato and Fusarium oxysporum. *Molecular Plant Pathology,* **11,** 309-314.

Taylor, A., Vagany, V., Barbara, D. J., Thomas, B., Pink, D. A. C., Jones, J. E. and Clarkson, J. P. (2013) Identification of differential resistance to six Fusarium oxysporum f. sp cepae isolates in commercial onion cultivars through the development of a rapid seedling assay. *Plant Pathology,* **62,** 103-111.

Taylor, A., Vagany, V., Jackson, A. C., Harrison, R. J., Rainoni, A. and Clarkson, J. P. (2016) Identification of pathogenicity-related genes in Fusarium oxysporum f. sp cepae. *Molecular Plant Pathology,* **17,** 1032-1047.

Testa, A. C., Hane, J. K., Ellwood, S. R. and Oliver, R. P. (2015) CodingQuarry: highly accurate hidden Markov model gene prediction in fungal genomes using RNA-seq transcripts. *BMC Genomics,* **16,** 170.

Thatcher, L. F., Gardiner, D. M., Kazan, K. and Manners, J. M. (2012) A highly conserved effector in Fusarium oxysporum is required for full virulence on Arabidopsis. *Molecular Plant-Microbe Interactions,* **25,** 180-190.

Thatcher, L. F., Williams, A. H., Garg, G., Buck, S. A. G. and Singh, K. B. (2016) Transcriptome analysis of the fungal pathogen Fusarium oxysporum f. sp medicaginis during colonisation of resistant and susceptible Medicago truncatula hosts identifies differential pathogenicity profiles and novel candidate effectors. *BMC Genomics,* **17,** 860.

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J. and Pachter, L. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology,* **28,** 511-515.

Tulbek, M. C., Lam, R. S. H., Wang, Y., Asavajaru, P. and Lam, A. (2017) Chapter 9 - Pea: a sustainable vegetable protein crop. In: *Sustainable Protein Sources.* (Nadathur, S. R., Wanasundara, J. P. D. and Scanlin, L., eds.). San Diego: Academic Press,

Tyson, J. R., O'Neil, N. J., Jain, M., Olsen, H. E., Hieter, P. and Snutch, T. P. (2018) MinION-based long-read sequencing and assembly extends the Caenorhabditis elegans reference genome. *Genome Research,* **28,** 266-274.

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R. and Leunissen, J. A. M. (2007) Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research,* **35,** W71-W74.

Urban, M., Cuzick, A., Rutherford, K., Irvine, A., Pedro, H., Pant, R., Sadanadan, V., Khamari, L., Billal, S., Mohanty, S. and Hammond-Kosack, K. E. (2017) PHI-base: a new interface and further additions for the multi-species pathogen-host interactions database. *Nucleic Acids Research,* **45,** D604-D610.

Vágány, V. (2012) Characterisation of Fusarium pathogens in the UK. PhD Thesis, School of Life Sciences, University of Warwick.

van Dam, P., de Sain, M., ter Horst, A., van der Gragt, M. and Rep, M. (2018) Use of comparative genomics-based markers for discrimination of host specificity in Fusarium oxysporum. *Applied and Environmental Microbiology,* **84,** e01868-01817.

van Dam, P., Fokkens, L., Ayukawa, Y., van der Gragt, M., ter Horst, A., Brankovics, B., Houterman, P. M., Arie, T. and Rep, M. (2017) A mobile pathogenicity chromosome in Fusarium oxysporum for infection of multiple cucurbit species. *Scientific Reports,* **7,** 9042.

van Dam, P., Fokkens, L., Schmidt, S. M., Linmans, J. H. J., Kistler, H. C., Ma, L. J. and Rep, M. (2016) Effector profiles distinguish formae speciales of Fusarium oxysporum. *Environmental Microbiology,* **18,** 4087-4102.

van Dam, P. and Rep, M. (2017) The distribution of Miniature Impala Elements and SIX genes in the Fusarium genus is suggestive of horizontal gene transfer. *Journal of Molecular Evolution,* **85,** 14-25.

van der Does, H. C., Duyvesteijn, R. G. E., Goltstein, P. M., van Schie, C. C. N., Manders, E. M. M., Cornelissen, B. J. C. and Rep, M. (2008) Expression of effector gene SIX1 of Fusarium oxysporum requires living plant cells. *Fungal Genetics and Biology,* **45,** 1257-1264.

van der Does, H. C., Fokkens, L., Yang, A., Schmidt, S. M., Langereis, L., Lukasiewicz, J. M., Hughes, T. R. and Rep, M. (2016) Transcription factors encoded on core and accessory chromosomes of Fusarium oxysporum induce expression of effector genes. *Plos Genetics,* **12,** e1006401.

van Diepeningen, A. D., Al-Hatmi, A. M. S., Brankovics, B. and de Hoog, G. S. (2014) Taxonomy and clinical spectra of Fusarium species: where do we stand in 2014? *Current Clinical Microbiology Reports,* **1,** 10-18.

Vaser, R., Sović, I., Nagarajan, N. and Šikić, M. (2017) Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Research,* **27,** 737-746.

Vlaardingerbroek, I., Beerens, B., Schmidt, S. M., Cornelissen, B. J. C. and Rep, M. (2016) Dispensable chromosomes in Fusarium oxysporum f. sp lycopersici. *Molecular Plant Pathology,* **17,** 1455-1466.

Vocanson, A., Roger-Estrade, J., Boizard, H. and Jeuffroy, M. H. (2006) Effects of soil structure on pea (Pisum sativum L.) root development according to sowing date and cultivar. *Plant and Soil,* **281,** 121-135.

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q. D., Wortman, J., Young, S. K. and Earl, A. M. (2014) Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *Plos One,* **9,** e112963.

Weirather, J. L., de Cesare, M., Wang, Y., Piazza, P., Sebastiano, V., Wang, X.-J., Buck, D. and Au, K. F. (2017) Comprehensive comparison of Pacific Biosciences and Oxford Nanopore Technologies and their applications to transcriptome analysis. *F1000Research,* **6,** 100-100.

Whalley, W. M. (1984) A rapid test for assessing the reaction of pea cultivars to pathogenic isolates of Fusarium solani f. sp pisi. *Annals of Applied Biology,* **104 (Supplementary),** 118-119.

Williams, A. H., Sharma, M., Thatcher, L. F., Azam, S., Hane, J. K., Sperschneider, J., Kidd, B. N., Anderson, J. P., Ghosh, R., Garg, G., Lichtenzveig, J., Kistler, C. H., Shea, T., Young, S., Buck, S.-A. G., Kamphuis, L. G., Saxena, R., Pande, S., Ma, L.-J., Varshney, R. K. and Singh, K. B. (2016) Comparative genomics and prediction of conditionally dispensable sequences in legume–infecting Fusarium oxysporum formae speciales facilitates identification of candidate effectors. *BMC Genomics,* **17,** 191.

Wilson, D., Charoensawan, V., Kummerfeld, S. K. and Teichmann, S. A. (2008) DBDtaxonomically broad transcription factor predictions: new content and functionality. *Nucleic Acids Research,* **36,** D88-D92.

Xue, A. G. (2003) Biological control of pathogens causing root rot complex in field pea using Clonostachys rosea strain ACM941. *Phytopathology,* **93,** 329-335.

Yadav, S., Maurya, S. K., Anand, G., Dwivedi, R. and Yadav, D. (2017) Purification, characterization and retting of Crotolaria juncea fibres by an alkaline pectin lyase from Fusarium oxysporum MTCC 1755. *3 Biotech,* **7,** 136.

Yang, H., Shi, P. J., Liu, Y., Xia, W., Wang, X. Y., Cao, H. F., Ma, R., Luo, H. Y., Bai, Y. G. and Yao, B. (2017) Loop 3 of fungal endoglucanases of glycoside hydrolase family 12 modulates catalytic efficiency. *Applied and Environmental Microbiology,* **83,** e03123-03116.

yes-peas! (2018) yes peas! Facts and Figures. http://www.peas.org (Accessed 2018).

Zhang, Z. N., Wu, Q. Y., Zhang, G. Z., Zhu, Y. Y., Murphy, R. W., Liu, Z. and Zou, C. G. (2015) Systematic analyses reveal uniqueness and origin of the CFEM domain in fungi. *Scientific Reports,* **5,** 13032.

Zitnick-Anderson, K., Simons, K. and Pasche, J. S. (2018) Detection and qPCR quantification of seven Fusarium species associated with the root rot complex in field pea. *Canadian Journal of Plant Pathology,* **40,** 261-271.

Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research,* **31,** 3406-3415.

Appendix 2

Table A 2.1 Putative *Fusarium* isolates obtained from the Processors and Growers Research Organisation (PGRO), identified by sequencing of the *translation elongation factor 1α* (*TEF*) gene.

Isolate	Identity (*TEF*)	Source	Location
PG1	*F. oxysporum*	Shropshire	Shropshire
PG2	*F. oxysporum*	Shropshire	Shropshire
PG3	*F. oxysporum*	Shropshire	Shropshire
PG4	*F. oxysporum*	Shropshire	Shropshire
PG5	*F. poae*	Seed	Unknown
PG6	*F. poae*	Seed	Unknown
PG7	*F. poae*	Seed	Unknown
PG8	*F. avenaceum*	Seed	Unknown
PG9	*F. avenaceum*	Seed	Unknown
PG10	*F. avenaceum*	Seed	Unknown
PG11	*F. avenaceum*	Seed testing from plants	Unknown
PG12	*F. culmorum*	Seed testing from plants	Unknown
PG13	*F. solani*	Crop clinic 2012	East Anglia
PG14	*F. solani*	Crop clinic 2012	East Anglia
PG15	*F. oxysporum*	Crop clinic 2012	East Anglia
PG16	*F. oxysporum*	Crop clinic 2012	East Anglia
PG17	*F. oxysporum*	Crop clinic 2012	East Anglia
PG18	*F. oxysporum*	Crop clinic 2012	East Anglia
PG19	*F. oxysporum*	Crop clinic 2012	East Anglia
PG20	*F. oxysporum*	Crop clinic 2012	East Anglia
PG21	*F. oxysporum*	Crop clinic 2012	East Anglia
PG22	*F. redolens*	abstress plot 1	Thorney, Peterborough
PG23	*F. redolens*	abstress plot 1	Thorney, Peterborough
PG24	*F. redolens*	abstress plot 2	Thorney, Peterborough
PG25	*F. redolens*	abstress plot 2	Thorney, Peterborough
PG26	*F. redolens*	abstress plot 5	Thorney, Peterborough
PG27	*F. avenaceum*	abstress plot 6	Thorney, Peterborough
PG28	*F. redolens*	abstress plot 6	Thorney, Peterborough
PG29	*F. redolens*	abstress plot 10	Thorney, Peterborough
PG30	*F. redolens*	abstress plot 10	Thorney, Peterborough
PG31	*F. culmorum*	abstress plot 12	Thorney, Peterborough
PG32	*F. redolens*	abstress plot 13	Thorney, Peterborough
PG33	*F. culmorum*	abstress plot 16	Thorney, Peterborough
PG34	*F. redolens*	abstress plot 18	Thorney, Peterborough
PG35	*F. redolens*	abstress plot 19	Thorney, Peterborough
PG36	*F. redolens*	abstress plot 25	Thorney, Peterborough
PG37	*F. culmorum*	abstress plot 27	Thorney, Peterborough
PG38	*F. culmorum*	abstress plot 30	Thorney, Peterborough
PG39	*F. redolens*	abstress plot 30	Thorney, Peterborough
PG40	*F. redolens*	abstress plot 38	Thorney, Peterborough
PG41	*F. redolens*	abstress plot 38	Thorney, Peterborough
PG42	*F. redolens*	abstress plot 39	Thorney, Peterborough
PG43	*F. redolens*	abstress plot 39	Thorney, Peterborough
PG44	*F. culmorum*	abstress plot 39	Thorney, Peterborough
PG45	*F. redolens*	abstress plot 39	Thorney, Peterborough
PG46	*F. redolens*	abstress plot 61	Thorney, Peterborough
PG47	*F. redolens*	abstress plot 48	Thorney, Peterborough
PG48	*F. redolens*	abstress plot 62	Thorney, Peterborough

Isolate	Identity (*TEF*)	Source	Location
PG49	*F. avenaceum*	abstress plot 64	Thorney, Peterborough
PG50	*F. redolens*	abstress plot 65	Thorney, Peterborough
PG51	*F. redolens*	abstress plot 69	Thorney, Peterborough
PG52	*F. culmorum*	abstress plot 69	Thorney, Peterborough
PG53	*F. culmorum*	abstress plot 69	Thorney, Peterborough
PG54	*F. culmorum*	abstress plot 69	Thorney, Peterborough
PG55	*F. culmorum*	abstress plot 76	Thorney, Peterborough
PG56	*F. redolens*	abstress plot 76	Thorney, Peterborough
PG57	*F. oxysporum*	A14/26	East Anglia
PG58	*F. oxysporum*	A14/27	East Anglia
PG59	*F. oxysporum*	A14/27	East Anglia
PG60	*F. oxysporum*	A14/43	Northhants
PG61	*F. oxysporum*	A14/43	Northhants
PG62	*F. oxysporum*	A14/43	Northhants
PG63	*F. oxysporum*	A14/37	Warwickshire
PG64	*F. oxysporum*	A14/36	Worcestershire
PG65	*F. oxysporum*	A14/38	Northhants
PG66	*Pythium*	A14/42	Cambridgeshire
PG67	*F. oxysporum*	A14/45	Yorkshire
PG68	*F. avenaceum*	A14/46	Yorkshire
PG69	*F. redolens*	A14/46	Yorkshire
PG70	*F. avenaceum*	A14/46	Yorkshire
PG71	*F. avenaceum*	A14/47	Yorkshire
PG72	*F. oxysporum*	A14/51	Oxfordshire
PG73	*F. oxysporum*	A14/51	Oxfordshire
PG74	*F. oxysporum*	A14/51	Oxfordshire
PG75	*F. culmorum*	A14/51	Oxfordshire
PG76	*F. oxysporum*	A14/51	Oxfordshire
PG77	*F. oxysporum*	A14/51	Oxfordshire
PG78	*F. avenaceum*	Beans @Tilney	Tilney
PG79	*F. oxysporum*	garlic and Awais trial	Unknown
PG80	*F. poae*	used in awais trial	Unknown
PG81	*F. redolens*	abstress plot 61	Thorney, Peterborough
PG82	*F. redolens*	abstress plot 48	Thorney, Peterborough
PG83	*F. redolens*	abstress plot 65	Thorney, Peterborough
PG84	*F. redolens*	abstress plot 76	Thorney, Peterborough
PG85	*F. oxysporum*	A14/36	Worcestershire
PG86	*F. oxysporum*	A14/38	Northhants
PG87	*F. redolens*	A14/46	Yorkshire
PG88	*Mucor hiemalis*	A14/57	Scotland

Table A 2.2 Molecular identification of putative *Fusarium* isolates from diseased pea roots and stems from UK fields, sampled during the 2015 and 2016 growing seasons. The morph. category was assigned based on the morphology of mycelium on PDA plates. Isolate identity was determined by sequencing of the *translation elongation factor 1α* (*TEF*) gene.

Isolate	Identity (*TEF*)	Source	Location	Sampling year	Isolated from	Pea Cultivar	Morph. category
PG089	*F. solani*	Field 127	Yorkshire	2015	Root	Oasis	6
PG091	*F. solani*	Field 127	Yorkshire	2015	Root	Oasis	5
PG094	*F. oxysporum*	Westfield 3	Yorkshire	2015	Root	Unknown	3
PG095	*F. redolens*	Westfield 3	Yorkshire	2015	Root	Unknown	10
PG100	*F. solani*	Field 141	Yorkshire	2015	Root	Oasis	1
PG101	*F. oxysporum*	Field 141	Yorkshire	2015	Root	Oasis	3
PG108	*F. oxysporum*	Field 116	Yorkshire	2015	Root	Oasis	3
PG110	*F. oxysporum*	Field 116	Yorkshire	2015	Root	Oasis	9
PG113	*F. oxysporum*	Field 133	Yorkshire	2015	Root	Oasis	3
PG114	*F. oxysporum*	Field 133	Yorkshire	2015	Root	Oasis	14
PG116	*F. oxysporum*	Field 119	Yorkshire	2015	Root	Oasis	9
PG118	*F. oxysporum*	Field 119	Yorkshire	2015	Root	Oasis	3
PG126	*F. redolens*	Westfield 1	Yorkshire	2015	Root	Unknown	10
PG128	*F. oxysporum*	Field 126	Yorkshire	2015	Root	Oasis	9
PG130	*F. avenaceum*	Field 130	Yorkshire	2015	Root	O893	7
PG132	*F. solani*	Field 132	Yorkshire	2015	Root	O893	6
PG134	*F. solani*	Field 120	Yorkshire	2015	Root	Oasis	6
PG137	*F. oxysporum*	Field 125	Yorkshire	2015	Root	Oasis	3
PG138	*F. solani*	Field 125	Yorkshire	2015	Root	Oasis	1
PG139	*F. oxysporum*	Field 118	Yorkshire	2015	Root	Oasis	9
PG140	*F. oxysporum*	Field 118	Yorkshire	2015	Root	Oasis	3
PG143	*F. solani*	Field 137	Yorkshire	2015	Root	Standana	1
PG146	*Trichoderma*	Field 137	Yorkshire	2015	Root	Standana	2
PG150	*F. solani*	Field 124	Yorkshire	2015	Root	Oasis	6
PG151	*F. solani*	Field 140	Yorkshire	2015	Root	Oasis	1
PG152	*F. oxysporum*	Field 140	Yorkshire	2015	Root	Oasis	3
PG154	*F. oxysporum*	Tibthorpe Birdseye	Yorkshire	2015	Root	Unknown	10
PG155	*F. redolens*	Tibthorpe Birdseye	Yorkshire	2015	Root	Unknown	10
PG160	*F. solani*	Field 128	Yorkshire	2015	Root	Oasis	6
PG161	*F. culmorum*	Field 128	Yorkshire	2015	Root	Oasis	4
PG166	*F. solani*	Field 138	Yorkshire	2015	Root	Oasis	5
PG167	*F. solani*	Field 123	Yorkshire	2015	Root	Oasis	1
PG169	*Clonostachys* or *Bionectria*	Field 123	Yorkshire	2015	Root	Oasis	8
PG172	*F. solani*	Field 121	Yorkshire	2015	Root	Oasis	5
PG176	*F. oxysporum*	Field 135	Yorkshire	2015	Root	Oasis	3
PG180	*F. oxysporum*	Field 122	Yorkshire	2015	Root	Oasis	2
PG181	*F. oxysporum*	Field 122	Yorkshire	2015	Root	Oasis	9
PG185	*F. culmorum*	Folly Farm	Scotland	2015	Root	Unknown	4
PG187	*F. culmorum*	BE8	Yorkshire	2015	Root	Swallow	4
PG189	*F. redolens*	BE8	Yorkshire	2015	Root	Swallow	10
PG191	*F. oxysporum*	Field 119 (2)	Yorkshire	2015	Root	Oasis	3
PG193	*F. solani*	Field 117	Yorkshire	2015	Root	Oasis	6
PG196	*F. oxysporum*	BE7	Yorkshire	2015	Root	Novella	3
PG197	*F. oxysporum*	BE7	Yorkshire	2015	Root	Novella	1
PG198	*F. redolens*	BE3	Yorkshire	2015	Root	Wagtail	10
PG199	*Trichoderma*	BE3	Yorkshire	2015	Root	Wagtail	15
PG201	*F. avenaceum*	Birdseye (118)	Yorkshire	2015	Root	Unknown	12
PG204	*F. oxysporum*	Field 131	Yorkshire	2015	Root	O893	3

Isolate	Identity (*TEF*)	Source	Location	Sampling year	Isolated from	Pea Cultivar	Morph. category
PG205	*F. redolens*	BE6	Yorkshire	2015	Root	Unknown	10
PG206	*F. redolens*	BE6	Yorkshire	2015	Root	Unknown	10
PG209	*F. avenaceum*	BE4	Yorkshire	2015	Root	Wagtail	7
PG211	*F. graminearum*	BE5	Yorkshire	2015	Root	Kite	4
PG213	*F. avenaceum*	BE5	Yorkshire	2015	Root	Kite	7
PG222	*F. oxysporum*	Field 130	Yorkshire	2015	Root	O893	3
PG224	*F. oxysporum*	Field 130	Yorkshire	2015	Root	O893	3
PG225	*F. oxysporum*	Field 130	Yorkshire	2015	Root	O893	3
PG227	*F. oxysporum*	Field 130	Yorkshire	2015	Root	O893	9
PG228	*F. oxysporum*	Field 461 1A	Suffolk	2015	Root	Naches	15
PG231	*F. oxysporum*	Field 461 12A	Suffolk	2015	Root	Naches	2
PG232	*F. oxysporum*	Field 461 12A	Suffolk	2015	Root	Naches	9
PG233	*F. oxysporum*	Field 461 2A	Suffolk	2015	Root	Naches	15
PG236	*F. solani*	Field 461 6A	Suffolk	2015	Root	Naches	6
PG237	*F. oxysporum*	Field 461 6A	Suffolk	2015	Root	Naches	3
PG239	*F. solani*	Field 461 5A	Suffolk	2015	Root	Naches	1
PG242	*F. oxysporum*	Field 461 4A	Suffolk	2015	Root	Naches	14
PG246	*F. redolens*	Field 461 8B	Suffolk	2015	Root	Naches	10
PG247	*F. oxysporum*	Field 461 8B	Suffolk	2015	Root	Naches	15
PG248	*F. oxysporum*	Field 121	Yorkshire	2015	Root	Oasis	3
PG250	*F. oxysporum*	Westfield 1	Yorkshire	2015	Root	Unknown	3
PG252	*F. oxysporum*	Field 461 10B	Suffolk	2015	Root	Naches	3
PG253	*F. solani*	Field 461 11A	Suffolk	2015	Root	Naches	1
PG254	*F. oxysporum*	Field 461 11A	Suffolk	2015	Root	Naches	15
PG256	*F. redolens*	Field 447 (1)	Suffolk	2015	Root	Standana	14
PG257	*F. lacertarum*	Field 447 (1)	Suffolk	2015	Root	Standana	16
PG259	*F. oxysporum*	Field 450	Suffolk	2015	Root	Terrain	3
PG260	*F. solani*	Field 461 14B	Suffolk	2015	Root	Naches	2
PG261	*F. oxysporum*	Field 461 14B	Suffolk	2015	Root	Naches	3
PG262	*Trichoderma*	Field 138	Yorkshire	2015	Root	Oasis	16
PG264	*F. oxysporum*	Field 452	Suffolk	2015	Root	Terrain	3
PG265	*F. oxysporum*	Field 452	Suffolk	2015	Root	Terrain	3
PG268	*F. oxysporum*	Field 443	Suffolk	2015	Root	Naches	3
PG270	*F. oxysporum*	Field 461 3B	Suffolk	2015	Root	Naches	15
PG271	*F. solani*	Field 461 3B	Suffolk	2015	Root	Naches	6
PG272	*F. solani*	Field 461 7B	Suffolk	2015	Root	Naches	1
PG273	*F. solani*	Birdseye (118)	Yorkshire	2015	Root	Unknown	6
PG275	*F. culmorum*	Field 458 (2)	Suffolk	2015	Root	Naches	4
PG277	*F. oxysporum*	Field 458 (2)	Suffolk	2015	Root	Naches	3
PG281	*F. oxysporum*	Field 461 9B	Suffolk	2015	Root	Naches	15
PG282	*F. culmorum*	Field 461 9B	Suffolk	2015	Root	Naches	4
PG283	*F. redolens*	Field 461 14A	Suffolk	2015	Root	Naches	10
PG284	*F. oxysporum*	Field 461 14A	Suffolk	2015	Root	Naches	15
PG286	*F. solani*	Field 461 4A	Suffolk	2015	Root	Naches	1
PG288	*F. oxysporum*	Field 445	Suffolk	2015	Root	Naches	3
PG290	*F. culmorum*	Field 445	Suffolk	2015	Root	Naches	4
PG291	*F. solani*	Field 461 6B	Suffolk	2015	Root	Naches	6
PG292	*F. oxysporum*	Field 461 6B	Suffolk	2015	Root	Naches	13
PG293	*F. oxysporum*	Field 454	Suffolk	2015	Root	Naches	14
PG294	*F. oxysporum*	Field 454	Suffolk	2015	Root	Naches	3
PG296	*F. oxysporum*	Field 119 (2)	Yorkshire	2015	Root	Oasis	3
PG300	*F. culmorum*	Field 451	Suffolk	2015	Root	Terrain	4
PG301	*F. oxysporum*	Field 451	Suffolk	2015	Root	Terrain	9
PG303	*F. oxysporum*	Field 461 4B	Suffolk	2015	Root	Naches	15
PG304	*F. oxysporum*	Field 461 4B	Suffolk	2015	Root	Naches	3
PG305	*F. solani*	Field 461 5B	Suffolk	2015	Root	Naches	5
PG306	*F. culmorum*	Field 461 5B	Suffolk	2015	Root	Naches	4
PG308	*F. oxysporum*	Field 459 Right	Suffolk	2015	Root	Naches	13
PG309	*F. oxysporum*	Field 459 Right	Suffolk	2015	Root	Naches	2
PG312	*F. oxysporum*	Field 461 15A	Suffolk	2015	Root	Naches	2
PG313	*F. oxysporum*	Field 461 15A	Suffolk	2015	Root	Naches	13
PG314	*F. avenaceum*	Field 461 11B	Suffolk	2015	Root	Naches	12
PG316	*F. oxysporum*	Field 461 11B	Suffolk	2015	Root	Naches	14
PG319	*F. solani*	Field 438	Suffolk	2015	Root	Butana	1

Isolate	Identity (*TEF*)	Source	Location	Sampling year	Isolated from	Pea Cultivar	Morph. category
PG321	*F. oxysporum*	Field 438	Suffolk	2015	Root	Butana	2
PG323	*F. solani*	Field 130	Yorkshire	2015	Root	O893	1
PG327	*F. oxysporum*	Field 461 13A	Suffolk	2015	Root	Naches	3
PG329	*F. culmorum*	Field 461 13A	Suffolk	2015	Root	Naches	4
PG331	*F. culmorum*	Field 461 7B	Suffolk	2015	Root	Naches	4
PG333	*F. oxysporum*	Field 461 9A	Suffolk	2015	Root	Naches	15
PG335	*F. oxysporum*	Field 461 9A	Suffolk	2015	Root	Naches	9
PG336	*F. oxysporum*	BE1	Yorkshire	2015	Root	Kite	3
PG337	*F. oxysporum*	BE1	Yorkshire	2015	Root	Kite	9
PG340	*F. oxysporum*	BE2	Yorkshire	2015	Root	Swallow	2
PG342	*F. oxysporum*	BE2	Yorkshire	2015	Root	Swallow	3
PG345	*F. oxysporum*	BE4	Yorkshire	2015	Root	Wagtail	9
PG347	*F. culmorum*	Field 449	Suffolk	2015	Root	Terrain	4
PG349	*F. avenaceum*	Field 449	Suffolk	2015	Root	Terrain	7
PG350	*F. redolens*	Field 448	Suffolk	2015	Root	Naches	3
PG351	*F. oxysporum*	Field 448	Suffolk	2015	Root	Naches	14
PG354	*F. oxysporum*	Field 458 (1)	Suffolk	2015	Root	Naches	2
PG355	*F. oxysporum*	Field 458 (1)	Suffolk	2015	Root	Naches	3
PG357	*F. redolens*	Field 461 13B	Suffolk	2015	Root	Naches	10
PG358	*F. oxysporum*	Field 461 13B	Suffolk	2015	Root	Naches	3
PG359	*F. oxysporum*	Field 444	Suffolk	2015	Root	EX	14
PG362	*F. solani*	Field 444	Suffolk	2015	Root	EX	1
PG364	*F. solani*	Field 117	Yorkshire	2015	Root	Oasis	1
PG365	*F. solani*	Field 131	Yorkshire	2015	Root	O893	5
PG369	*F. oxysporum*	Field 447 (2)	Suffolk	2015	Root	Standana	13
PG370	*F. redolens*	Field 447 (2)	Suffolk	2015	Root	Standana	10
PG373	*F. oxysporum*	Field 461 10A	Suffolk	2015	Root	Naches	15
PG374	*F. oxysporum*	Field 461 10A	Suffolk	2015	Root	Naches	14
PG375	*F. oxysporum*	Field 461 7A	Suffolk	2015	Root	Naches	15
PG378	*F. oxysporum*	Field 461 2B	Suffolk	2015	Root	Naches	3
PG379	*F. oxysporum*	Field 461 2B	Suffolk	2015	Root	Naches	9
PG380	*F. oxysporum*	Field 461 3A	Suffolk	2015	Root	Naches	15
PG381	*F. oxysporum*	Field 461 3A	Suffolk	2015	Root	Naches	3
PG383	*F. oxysporum*	Field 461 1B	Suffolk	2015	Root	Naches	3
PG384	*F. solani*	Field 461 1B	Suffolk	2015	Root	Naches	1
PG386	*F. lacertarum*	Field 462	Suffolk	2015	Root	Odet	2
PG388	*F. lacertarum*	Field 462	Suffolk	2015	Root	Odet	16
PG389	*F. oxysporum*	Field 461 12B	Suffolk	2015	Root	Naches	2
PG390	*F. oxysporum*	Field 461 12B	Suffolk	2015	Root	Naches	3
PG391	*F. oxysporum*	Field 459 Left	Suffolk	2015	Root	Naches	3
PG393	*F. oxysporum*	Field 459 Left	Suffolk	2015	Root	Naches	13
PG394	*F. avenaceum*	Field 443	Suffolk	2015	Root	Naches	12
PG395	*F. oxysporum*	Field 450	Suffolk	2015	Root	Terrain	14
PG399	*F. solani*	Field 461 8A	Suffolk	2015	Root	Naches	1
PG401	*F. oxysporum*	Field 461 8A	Suffolk	2015	Root	Naches	10
PG418	*F. graminearum*	OAS71	Lincolnshire	2016	Stem	Oasis	4
PG423	*F. solani*	OAS81	Lincolnshire	2016	Stem	Oasis	6
PG439	*F. solani*	Field 102	Yorkshire	2016	Stem	Grundy	6
PG442	*F. solani*	OAS73	Lincolnshire	2016	Stem	Oasis	1
PG454	*F. solani*	OAS68	Lincolnshire	2016	Stem	Oasis	2
PG457	*F. avenaceum*	Field 100	Yorkshire	2016	Stem	Ashton	7
PG458	*F. redolens*	OAS79	Lincolnshire	2016	Stem	Oasis	10
PG459	*F. redolens*	OAS79	Lincolnshire	2016	Stem	Oasis	10
PG460	*F. solani*	OAS79	Lincolnshire	2016	Stem	Oasis	6
PG461	*F. solani*	OAS73	Lincolnshire	2016	Stem	Oasis	1
PG462	*F. solani*	OAS73	Lincolnshire	2016	Stem	Oasis	1
PG463	*F. solani*	OAS73	Lincolnshire	2016	Stem	Oasis	1
PG464	*F. equiseti*	Field 108	Yorkshire	2016	Stem	Oasis	3
PG466	*F. solani*	Field 98	Yorkshire	2016	Stem	Ashton	1
PG467	*F. oxysporum*	Field 98	Yorkshire	2016	Stem	Ashton	2
PG469	*F. solani*	SES76	Lincolnshire	2016	Stem	Serge	1
PG470	*F. solani*	SES76	Lincolnshire	2016	Stem	Serge	6
PG472	*F. equiseti*	Field 105	Yorkshire	2016	Stem	Oasis	3
PG473	*F. avenaceum*	Field 105	Yorkshire	2016	Stem	Oasis	7

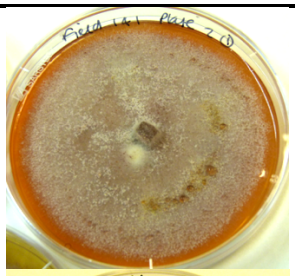
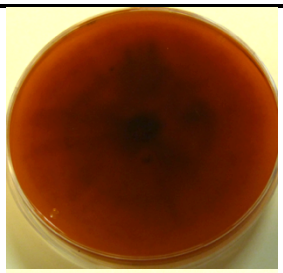
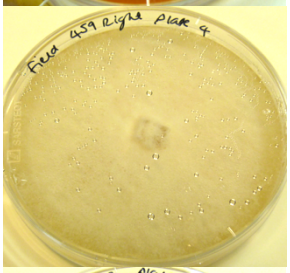
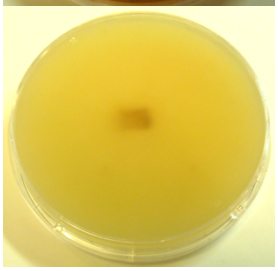
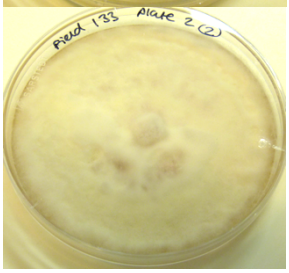
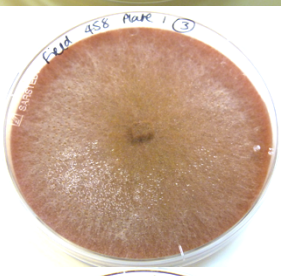
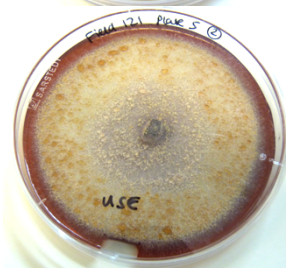
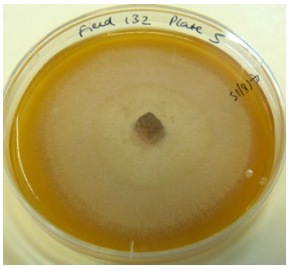
Isolate	Identity (*TEF*)	Source	Location	Sampling year	Isolated from	Pea Cultivar	Morph. category
PG474	*F. solani*	OAS71	Lincolnshire	2016	Stem	Oasis	6
PG475	*F. solani*	OAS71	Lincolnshire	2016	Stem	Oasis	1
PG476	*F. oxysporum*	OAS71	Lincolnshire	2016	Stem	Oasis	2
PG477	*F. solani*	Field 97	Yorkshire	2016	Stem	Ashton	3
PG478	*F. solani*	Field 97	Yorkshire	2016	Stem	Ashton	6
PG479	*F. solani*	Field 97	Yorkshire	2016	Stem	Ashton	6
PG480	*F. oxysporum*	Field 97	Yorkshire	2016	Stem	Ashton	2
PG481	*F. avenaceum*	Field 97	Yorkshire	2016	Stem	Ashton	7
PG482	*F. redolens*	OAS81	Lincolnshire	2016	Stem	Oasis	10
PG483	*F. solani*	Field 96	Yorkshire	2016	Stem	Ashton	1
PG484	*F. solani*	OAS72	Lincolnshire	2016	Stem	Oasis	1
PG485	*F. redolens*	OAS72	Lincolnshire	2016	Stem	Oasis	10
PG487	*F. redolens*	OAS78	Lincolnshire	2016	Stem	Oasis	10
PG490	*F. solani*	Field 97	Yorkshire	2016	Stem	Ashton	6
PG492	*F. solani*	Field 99	Yorkshire	2016	Stem	Ashton	6
PG493	*F. solani*	Field 99	Yorkshire	2016	Stem	Ashton	3
PG494	*F. oxysporum*	Field 99	Yorkshire	2016	Stem	Ashton	2
PG495	*F. solani*	Field 98	Yorkshire	2016	Stem	Ashton	1
PG496	*F. solani*	OAS81	Lincolnshire	2016	Stem	Oasis	6
PG497	*F. solani*	OAS73	Lincolnshire	2016	Stem	Oasis	3
PG499	*F. solani*	OAS78	Lincolnshire	2016	Stem	Oasis	2
PG500	*F. solani*	OAS72	Lincolnshire	2016	Stem	Oasis	1
PG501	*F. solani*	Field 98	Yorkshire	2016	Stem	Ashton	3
PG502	*F. solani*	OAS79	Lincolnshire	2016	Stem	Oasis	1
PG503	*F. solani*	OAS81	Lincolnshire	2016	Stem	Oasis	2
PG504	*F. solani*	OAS81	Lincolnshire	2016	Stem	Oasis	1

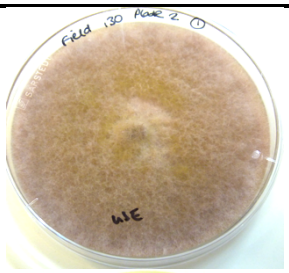
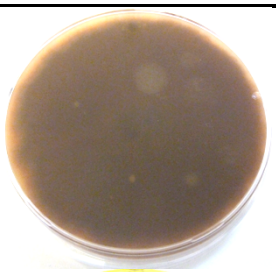
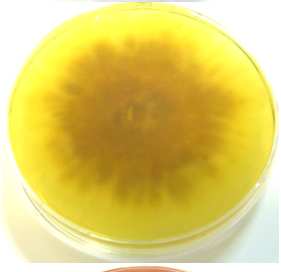
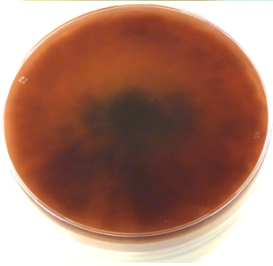
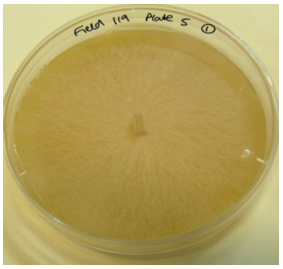
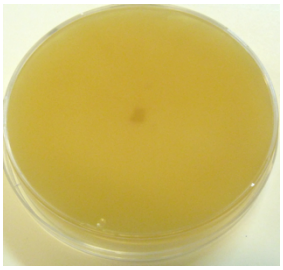
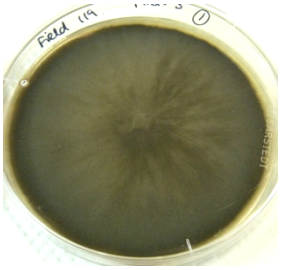
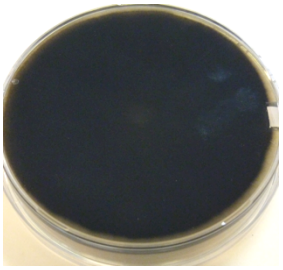
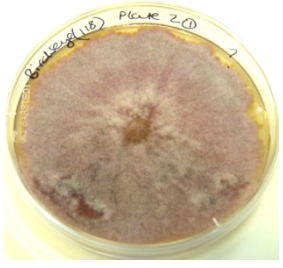
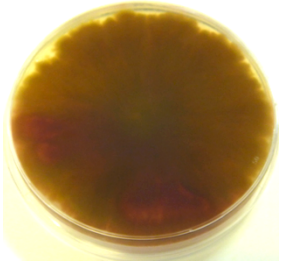
Table A 2.3 Molecular identification of putative *Fusarium oxysporum* f. sp. *pisi* (FOP) isolates from the UK and overseas. The original species/*forma specialis* and race type from the sender is listed. Identification was carried out based on sequencing of the *translation elongation factor 1α* (*TEF*) gene and a new race designated by *TEF* phylogeny and presence/absence of *Secreted In Xylem (SIX*) genes.

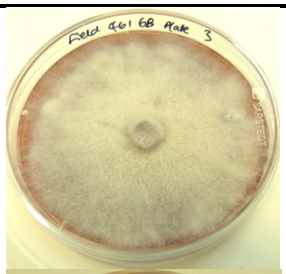
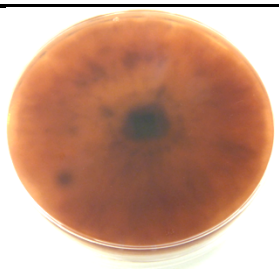
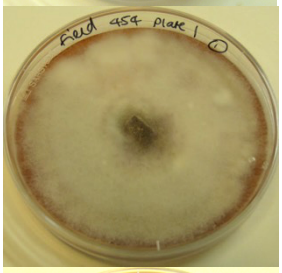
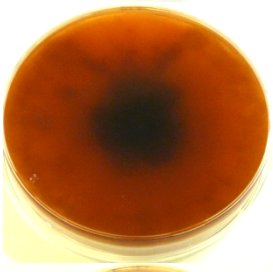
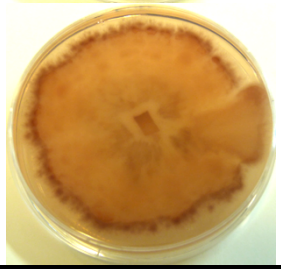
Isolate name	Species ID from sender	Orig. race	Country of origin	Contact name	Identified species (*TEF*)	Race (*TEF/SIX*)
CBS183.35	*F. oxysporum* f. sp. *pisi*	1	CBS culture collection	CBS	*F. oxysporum*	1
CBS170.30	*F. oxysporum* f. sp. *pisi*	1	CBS culture collection	CBS	*F. oxysporum*	1
CBS260.51	*F. oxysporum* f. sp. *pisi*	2	CBS culture collection	CBS	*F. oxysporum*	1
PDA3b	*F. oxysporum* f. sp. *pisi*	2	USA	Dr R. J McGee	*F. oxysporum*	5
F79	*F. oxysporum* f. sp. *pisi*	1	USA	Dr R. J McGee	*F. oxysporum*	1
Fw-09-E	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	RR
F236	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	2
F232	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	2
F30	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	2
F81	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	2
Fw-08-04	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	RR
F42a	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	5
Fw-08-02	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	RR
F16	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	2
F35	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	2
F31	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	2
Fw-08-03	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	RR
F235	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	2
F234	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	2
F231	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	2
F40	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	5
Fw-09-A	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. solani*	
F233	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	2
F237	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	2
Fw-09-D	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	1

Isolate name	Species ID from sender	Orig. race	Country of origin	Contact name	Identified species (*TEF*)	Race (*TEF/SIX*)
Fw-09-C	*F. oxysporum* f. sp. *pisi*	2	USA	Dr L. Porter	*F. oxysporum*	1
FOC	*F. oxysporum* f. sp. *ciceris*	2	USA	Dr J. Pasche	*F. oxysporum*	
AC1.2	*F. oxysporum* f. sp. *pisi*	1	Algeria	Dr A. Merzoug	*F. oxysporum*	RR
AL1.2	*F. oxysporum* f. sp. *pisi*	1	Algeria	Dr A. Merzoug	*F. avenaceum*	
AC6	*F. oxysporum* f. sp. *pisi*	6	Algeria	Dr A. Merzoug	*F. oxysporum*	RR
AL2.1	*F. oxysporum* f. sp. *pisi*	2	Algeria	Dr A. Merzoug	*F. oxysporum*	RR
AC2.2	*F. oxysporum* f. sp. *pisi*	2	Algeria	Dr A. Merzoug	*F. avenaceum*	
R1B	*F. oxysporum* f. sp. *pisi*	1	Czech Republic (via PGRO)	Dr L. Herold (PGRO)	*F. redolens*	
R1A	*F. oxysporum* f. sp. *pisi*	1	Czech Republic (via PGRO)	Dr L. Herold (PGRO)	*F. redolens*	
R2	*F. oxysporum* f. sp. *pisi*	2	Czech Republic (via PGRO)	Dr L. Herold (PGRO)	*F. oxysporum*	5
FOP1	*F. oxysporum* f. sp. *pisi*	1	UK (historic Warwick HRI isolate)	Dr C. Linfield	*F. oxysporum*	1
FOP2	*F. oxysporum* f. sp. *pisi*	2	UK (historic Warwick HRI isolate)	Dr C. Linfield	*F. oxysporum*	2
FOP5	*F. oxysporum* f. sp. *pisi*	5	UK (historic Warwick HRI isolate)	Dr C. Linfield	*F. oxysporum*	5
FOP1 EMR	*F. oxysporum* f. sp. *pisi*	1	UK	NIAB-EMR	*F. oxysporum*	1

Table A 2.4 Morphology of putative *Fusarium* isolates sampled from diseased peas in UK fields, separated into 15 categories based on colour, aerial mycelium and growth patterns.

Category	Aerial growth	Underside of plate	Description
1			Orange/red base with white/purple/orange bitty mycelium
2			Cream base, thick white mycelium
3			Cream/pink base, white mycelium
4			Deep red/brown base with red/brown flat mycelium
5			Deep purple base, white/orange/yellow/purple mycelium
6			Yellow/orange base, orange circles, yellow/white flat mycelium

Category	Aerial growth	Underside of plate	Description
7			Brown base, brown/red mycelium
8			Bright yellow base, pink centre, White fluffy mycelium
9			Red base, white mycelium
10			Cream base, pink centre. White spindly mycelium
11			Non-*Fusarium* species. Black mycelium with green edges
12			Brown base, red/pink/brown mycelium

Category	Aerial growth	Underside of plate	Description
13			Cream/dark pink bitty base. Growth unequal from centre. White mycelium
14			Bright orange base (agar), dark centre. White mycelium. Purple around middle cube of agar
15			Cream/pink base. Pink circle around surround edge of mycelium. White mycelium (can be veiny pink underneath).

Appendix 3

Table A 3.1 Genes in FOP1 EMR (race 1) belonging to orthogroups also shared by F81 (race 2) and R2 (race 5), identified by orthology analysis in the three isolates of *Fusarium oxysporum* f. sp. *pisi* (FOP). Functional annotation of genes including whether it was predicted as secreted, within 2 kb of a mimp, identified by EffectorP, the shared orthogroup, the number of genes in the orthogroup in each isolate and any annotation from InterProScan was included.

Gene	Contig	Sec.	2 kb of a mimp	EffP	CAZY	Orthogroup	Orthogroup count (F81:R2:F1EMR)	Annotation
g17046	19	Yes	Yes	Yes		2436	2:3:2	
g45	1	Yes		Yes		12380	1:1:1	
g19443	30	Yes		Yes		12822	1:1:1	
g19442	30	Yes		Yes		12821	1:1:1	
g46	1	Yes		Yes		12381	1:1:1	
g18161	24	Yes		Yes		12017	1:1:2	
g19102	28	Yes		Yes		12017	1:1:2	
g17700	22	Yes		Yes		1123	3:3:6	
g21974	68	Yes		Yes		1123	3:3:6	
g17039	19	Yes		Yes		1123	3:3:6	
g8816	6	Yes				12565	1:1:1	
g16069	16	Yes				11965	1:1:2	
g16256	16	Yes				11965	1:1:2	
g21456	51		Yes	Yes		12034	1:1:2	
g19234	29		Yes	Yes		6658	1:2:2	N-terminal beta barrel domain
g17407	20			Yes	CAZY :CE10	12785	1:1:1	Alpha/Beta hydrolase fold
g8340	5			Yes		12554	1:1:1	
g19273	29			Yes		12817	1:1:1	DinB/YfiT-like putative metal-dependent hydrolase
g17137	19			Yes		12779	1:1:1	
g8084	5			Yes		12549	1:1:1	
g8281	5			Yes		12552	1:1:1	Membrane-bound protein
g6850	4			Yes		12523	1:1:1	
g2226	1			Yes		12438	1:1:1	
g3610	2			Yes		12471	1:1:1	
g9487	6			Yes		12577	1:1:1	Zinc knuckle
g3474	2			Yes		12467	1:1:1	Transferase synthase holo-acyl-carrier-protein
g7873	5			Yes		12546	1:1:1	
g20649	40			Yes		12848	1:1:1	
g14387	12			Yes		12709	1:1:1	
g13150	10			Yes		12662	1:1:1	
g14227	11			Yes		12705	1:1:1	Aldehyde dehydrogenase
g7178	4			Yes		12528	1:1:1	
g14751	12			Yes		12715	1:1:1	
g14764	12			Yes		12717	1:1:1	
g21614	55			Yes		12862	1:1:1	Alcohol dehydrogenase
g21620	55			Yes		12863	1:1:1	Peptidase S8/S53 domain
g16335	16			Yes		11308	1:1:2	Pyridoxal phosphate-dependent transferase
g21541	53			Yes		11428	1:1:2	Zinc finger, C2H2-like

Gene	Contig	Sec.	2 kb of a mimp	EffP	CAZY	Orthogroup	Orthogroup count (F81:R2:F1EMR)	Annotation
g2220	1			Yes		11428	1:1:2	Zinc finger, C2H2-like
g21480	52			Yes		11971	1:1:2	
g16205	16			Yes		11971	1:1:2	
g19282	29			Yes		12034	1:1:2	
g16979	19			Yes		10964	1:1:3	
g15124	13			Yes		10897	1:1:3	
g16274	16			Yes		10964	1:1:3	
g17183	19			Yes		10897	1:1:3	
g19086	28			Yes		10897	1:1:3	
g20883	43			Yes		11077	1:1:3	
g19873	33			Yes		3184	1:4:2	
g21594	55			Yes		3184	1:4:2	
g16183	16			Yes		995	10:1:2	
g17972	23			Yes		675	11:2:3	
g18923	27			Yes		675	11:2:3	
g20781	42			Yes		675	11:2:3	
g16250	16			Yes		11975	2:1:1	
g16059	16			Yes		11964	2:1:1	
g20373	38			Yes		12044	2:1:1	
g13504	10			Yes		4842	2:2:2	
g16132	16			Yes		4842	2:2:2	
g19896	33			Yes		2436	2:3:2	
g19412	30			Yes		11161	3:1:1	(SH3) domain profile
g2219	1			Yes		2584	3:1:3	
g21542	53			Yes		2584	3:1:3	
g21542	53			Yes		2584	3:1:3	
g16967	19			Yes		1123	3:3:6	
g16967	19			Yes		1123	3:3:6	
g17059	19			Yes		1123	3:3:6	
g8800	5			Yes		131	4:12:27	
g15654	15			Yes		131	4:12:27	MYND finger; Zinc finger
g2273	1			Yes		131	4:12:27	
g2274	1			Yes		131	4:12:27	MYND finger; Zinc finger
g18982	27			Yes		131	4:12:27	
g17975	23			Yes		2053	4:2:2	Putative diphthamide synthesis protein
g21598	55			Yes		2053	4:2:2	Putative diphthamide synthesis protein
g16301	16			Yes		1022	5:5:2	Short-chain dehydrogenase/reductase
g17384	20			Yes		1022	5:5:2	Short-chain dehydrogenase/reductase
g21506	52			Yes		695	8:2:6	
g17123	19			Yes		695	8:2:6	
g17881	22			Yes		695	8:2:6	
g19018	28			Yes		695	8:2:6	
g20944	44			Yes		695	8:2:6	
g21403	50			Yes		695	8:2:6	
g21444	51		Yes			11756	1:1:2	
g21449	51		Yes			1985	2:2:4	Cytochrome P450
g39	1				CAZY :CBM 18	12378	1:1:1	Chitin-binding chitinase precursor signal glycosidase hydrolase lectin
g18838	27				CAZY :GH10 9	11671	1:1:2	Oxidoreductase family, NAD-binding Rossmann fold
g17173	19				CAZY :GH10 9	11671	1:1:2	NAD(P)-binding domain
g14873	12				CAZY :GH3	12390	1:1:1	Glycosyl hydrolases family 3 active site

Table A 3.2 Genes in F81 (race 2) belonging to orthogroups also shared by FOP1 EMR (race 1) and R2 (race 5), identified by orthology analysis in the three isolates of *Fusarium oxysporum* f. sp. *pisi* (FOP). Functional annotation of genes including whether it was predicted as secreted, within 2 kb of a mimp, identified by EffectorP, the shared orthogroup, the number of genes in the orthogroup in each isolate and any annotation from InterProScan was included.

Gene	Contig	Sec.	2 kb of a mimp	EffP	CAZY	Orthogroup	Orthogroup count (F81:R2:F1EMR)	Annotation
g18880	35	Yes	Yes	Yes		1123	3:3:6	
g18891	35	Yes	Yes	Yes		1123	3:3:6	
g19747	46	Yes	Yes	Yes		12821	1:1:1	
g19265	40	Yes		Yes		12017	1:1:2	
g14418	15	Yes		Yes		12380	1:1:1	
g19746	46	Yes		Yes		12822	1:1:1	
g15746	17	Yes		Yes		1123	3:3:6	
g14417	15	Yes		Yes		12381	1:1:1	
g19323	40	Yes				11965	1:1:2	
g11159	10	Yes				12565	1:1:1	
g19210	39	Yes				12659	1:1:1	
g17085	22		Yes	Yes		12817	1:1:1	DinB/YfiT-like putative metal-dependent hydrolase
g16557	20			Yes		4842	2:2:2	
g2386	2			Yes		12554	1:1:1	
g16533	20			Yes		4842	2:2:2	
g17612	25			Yes		11308	1:1:2	Pyridoxal phosphate-dependent transferase
g2327	2			Yes		12552	1:1:1	
g3720	3			Yes		12516	1:1:1	Region of a membrane-bound protein
g2131	2			Yes		12549	1:1:1	
g3202	3			Yes		12523	1:1:1	
g11315	10			Yes		12862	1:1:1	Alcohol dehydrogenase GroES-like domain
g9171	8			Yes		12471	1:1:1	
g4254	4			Yes		2584	3:1:3	
g4254	4			Yes		2584	3:1:3	
g4246	4			Yes		12438	1:1:1	
g19856	48			Yes		2584	3:1:3	
g4253	4			Yes		11428	1:1:2	Zinc finger, C2H2-like
g19704	46			Yes		12797	1:1:1	
g6687	6			Yes		12577	1:1:1	Zinc knuckle
g16063	18			Yes		695	8:2:6	
g17383	23			Yes		695	8:2:6	
g17891	27			Yes		695	8:2:6	
g11257	10			Yes		675	11:2:3	
g17876	27			Yes		695	8:2:6	
g20290	61			Yes		695	8:2:6	
g15882	18			Yes		695	8:2:6	
g17345	23			Yes		695	8:2:6	
g18844	35			Yes		1022	5:5:2	Short-chain dehydrogenase/reductase dehydrogenase
g19302	40			Yes		2436	2:3:2	
g6285	5			Yes		675	11:2:3	
g9423	8			Yes		12467	1:1:1	Phosphopantetheinyl transferase superfamily
g9743	9			Yes		675	11:2:3	
g12160	11			Yes		675	11:2:3	
g17202	22			Yes		675	11:2:3	

Gene	Contig	Sec.	2 kb of a mimp	EffP	CAZY	Orthogroup	Orthogroup count (F81:R2:F1EMR)	Annotation
g17771	26			Yes		675	11:2:3	
g17976	27			Yes		675	11:2:3	
g17983	27			Yes		675	11:2:3	
g20461	72			Yes		675	11:2:3	
g20467	72			Yes		675	11:2:3	
g15757	17			Yes		11666	2:1:1	Region of a membrane-bound protein
g1921	2			Yes		12546	1:1:1	
g17060	22			Yes		2436	2:3:2	
g15897	18			Yes		1022	5:5:2	short chain dehydrogenase
g10261	9			Yes		12709	1:1:1	
g18642	33			Yes		11666	2:1:1	pyridoxal phosphate-dependent enzyme
g17073	22			Yes		12034	1:1:2	
g17914	27			Yes		11161	3:1:1	
g15899	18			Yes		1022	5:5:2	short chain dehydrogenase
g14423	15			Yes		12378	1:1:1	
g17584	25			Yes		1022	5:5:2	short chain dehydrogenase
g11345	10			Yes		3184	1:4:2	
g12253	11			Yes		12849	1:1:1	
g15705	17			Yes		11161	3:1:1	Src homology 3 (SH3) domain
g16765	20			Yes		12662	1:1:1	
g17378	23			Yes		10897	1:1:3	
g18807	35			Yes		695	8:2:6	
g19224	39			Yes		11971	1:1:2	
g6290	5			Yes		12705	1:1:1	Aldehyde dehydrogenase
g9903	9			Yes		12717	1:1:1	
g10388	9			Yes		11499	2:1:1	Enoyl-(Acyl carrier protein) reductase
g11321	10			Yes		12863	1:1:1	Peptidase S8/S53 domain
g11341	10			Yes		2053	4:2:2	Putative diphthamide synthesis protein
g16526	19			Yes		12528	1:1:1	
g17592	25			Yes		11946	1:1:2	Fungal specific transcription factor domain
g17767	26			Yes		2053	4:2:2	Putative diphthamide synthesis protein
g17866	26			Yes		12044	2:1:1	
g17916	27			Yes		11161	3:1:1	Src homology 3 (SH3) domain profile
g18593	32			Yes		12029	1:2:1	
g18654	33			Yes		675	11:2:3	
g19038	37			Yes		2053	4:2:2	Putative diphthamide synthesis protein
g19057	37			Yes		12044	2:1:1	
g19313	40			Yes		11964	2:1:1	
g19978	51			Yes		11964	2:1:1	
g20457	72			Yes		2053	4:2:2	Putative diphthamide synthesis protein
g17083	22		Yes			12818	1:1:1	S-adenosyl-L-methionine-dependent methyltransferase
g17082	22		Yes			1985	2:2:4	Cytochrome P450
g17081	22		Yes			12033	1:1:2	NmrA-like family
g17084	22		Yes			1985	2:2:4	Cytochrome P450
g19526	43		Yes			12422	1:1:1	
g15653	17				CE10	12785	1:1:1	Alpha/Beta hydrolase fold
g15932	18				GH109	11671	1:1:2	NAD(P)-binding domain
g9805	9				GH3	12390	1:1:1	Glycoside hydrolase
g18070	28					4427	2:1:3	Galactose-binding domain

Table A 3.3 Genes in R2 (race 5) belonging to orthogroups also shared by FOP1 EMR (race 1) and F81 (race 2), identified by orthology analysis in the three isolates of *Fusarium oxysporum* f. sp. *pisi* (FOP). Functional annotation of genes including whether it was predicted as secreted, within 2 kb of a mimp, identified by EffectorP, the shared orthogroup, the number of genes in the orthogroup in each isolate and any annotation from InterProScan was included.

Gene	Contig	Sec.	2 kb of a mimp	EffP	CAZY	Orthogroup	Orthogroup count (F81:R2:F1EMR)	Annotation
g17990	20	Yes	Yes	Yes		1123	3:3:6	
g18001	20	Yes	Yes	Yes		1123	3:3:6	
g18184	21	Yes	Yes	Yes		12821	1:1:1	
g17944	20	Yes		Yes		12017	1:1:2	
g18183	21	Yes		Yes		12822	1:1:1	
g3956	3	Yes		Yes		12380	1:1:1	
g3955	3	Yes		Yes		12381	1:1:1	
g17971	20	Yes		Yes		1123	3:3:6	
g19237	33	Yes				12659	1:1:1	
g6934	4	Yes				12565	1:1:1	
g17893	20	Yes				11965	1:1:2	
g17484	18		Yes	Yes		12817	1:1:1	DinB/YfiT-like putative metal-dependent hydrolase
g16250	14		Yes	Yes		12422	1:1:1	
g8483	6			Yes		12554	1:1:1	Region of a membrane-bound protein
g16875	16			Yes		4842	2:2:2	
g15887	14			Yes		4842	2:2:2	
g17284	17			Yes		12779	1:1:1	
g1097	1			Yes		12471	1:1:1	
g8739	6			Yes		12549	1:1:1	
g15354	13			Yes		12847	1:1:1	
g8545	6			Yes		12552	1:1:1	
g3477	2			Yes		12516	1:1:1	Region of a membrane-bound protein
g8239	5			Yes		2584	3:1:3	
g15362	13			Yes		11967	1:1:2	
g2953	2			Yes		12523	1:1:1	
g8233	5			Yes		12438	1:1:1	
g15769	13			Yes		695	8:2:6	
g15592	13			Yes		695	8:2:6	
g17960	20			Yes		11666	2:1:1	Cys/Met metabolism, pyridoxal phosphate-dependent enzyme
g18121	21			Yes		12797	1:1:1	
g8238	5			Yes		11428	1:1:2	Zinc finger, C2H2-like
g6276	4			Yes		12577	1:1:1	Zinc knuckle
g8950	6			Yes		12546	1:1:1	
g15680	13			Yes		675	11:2:3	
g17959	20			Yes		9733	1:2:2	
g1351	1			Yes		12467	1:1:1	Phosphopantetheine-protein transferase domain transferase superfamily
g10823	8			Yes		12709	1:1:1	
g15407	13			Yes		2436	2:3:2	
g12770	10			Yes		3184	1:4:2	
g17080	16			Yes		10897	1:1:3	
g18449	23			Yes		3184	1:4:2	
g18721	26			Yes		3184	1:4:2	

Gene	Contig	Sec.	2 kb of a mimp	EffP	CAZY	Orthogroup	Orthogroup count (F81:R2:F1EMR)	Annotation
g11163	8			Yes		12715	1:1:1	
g17462	18			Yes		2436	2:3:2	
g17914	20			Yes		2436	2:3:2	
g15744	13			Yes		1022	5:5:2	Short-chain dehydrogenase/reductase
g17306	17		46	Yes		1022	5:5:2	Short-chain dehydrogenase/reductase
g15422	13			Yes		1022	5:5:2	Short-chain dehydrogenase/reductase
g15176	12			Yes		12705	1:1:1	Aldehyde dehydrogenase
g17303	17			Yes		1022	5:5:2	Short-chain dehydrogenase/reductase
g19030	30			Yes		3184	1:4:2	
g15730	13			Yes		1022	5:5:2	Short-chain dehydrogenase/reductase
g2633	2			Yes		12528	1:1:1	Membrane-bound protein
g17473	18			Yes		12034	1:1:2	
g19251	33			Yes		11971	1:1:2	
g11179	8			Yes		12717	1:1:1	
g12753	9			Yes		12443	1:1:1	
g15754	13			Yes		11946	1:1:2	
g16034	14			Yes		12662	1:1:1	
g17356	17			Yes		12044	2:1:1	
g18027	20			Yes		675	11:2:3	
g18222	21			Yes		11964	2:1:1	
g18724	26			Yes		2053	4:2:2	Putative diphthamide synthesis protein
g18744	26			Yes		12862	1:1:1	Aldehyde dehydrogenase
g18752	26			Yes		12863	1:1:1	Peptidase S8/S53 domain
g18876	27			Yes		2053	4:2:2	Putative diphthamide synthesis protein
g17482	18		Yes			12818	1:1:1	S-adenosyl-L-methionine-dependent methyltransferase
g17481	18		Yes			1985	2:2:4	Cytochrome P450
g17480	18		Yes			12033	1:1:2	NAD(P)-binding domain
g17483	18		Yes			1985	2:2:4	Cytochrome P450
g3962	3				CBM18	12378	1:1:1	Chitin recognition or binding domain signature
g15470	13				CE10	12785	1:1:1	Alpha/Beta hydrolase fold
g17271	17				GH109	11671	1:1:2	NAD(P)-binding domain
g11268	8				GH3	12390	1:1:1	Fibronectin type III-like domain

Appendix 4

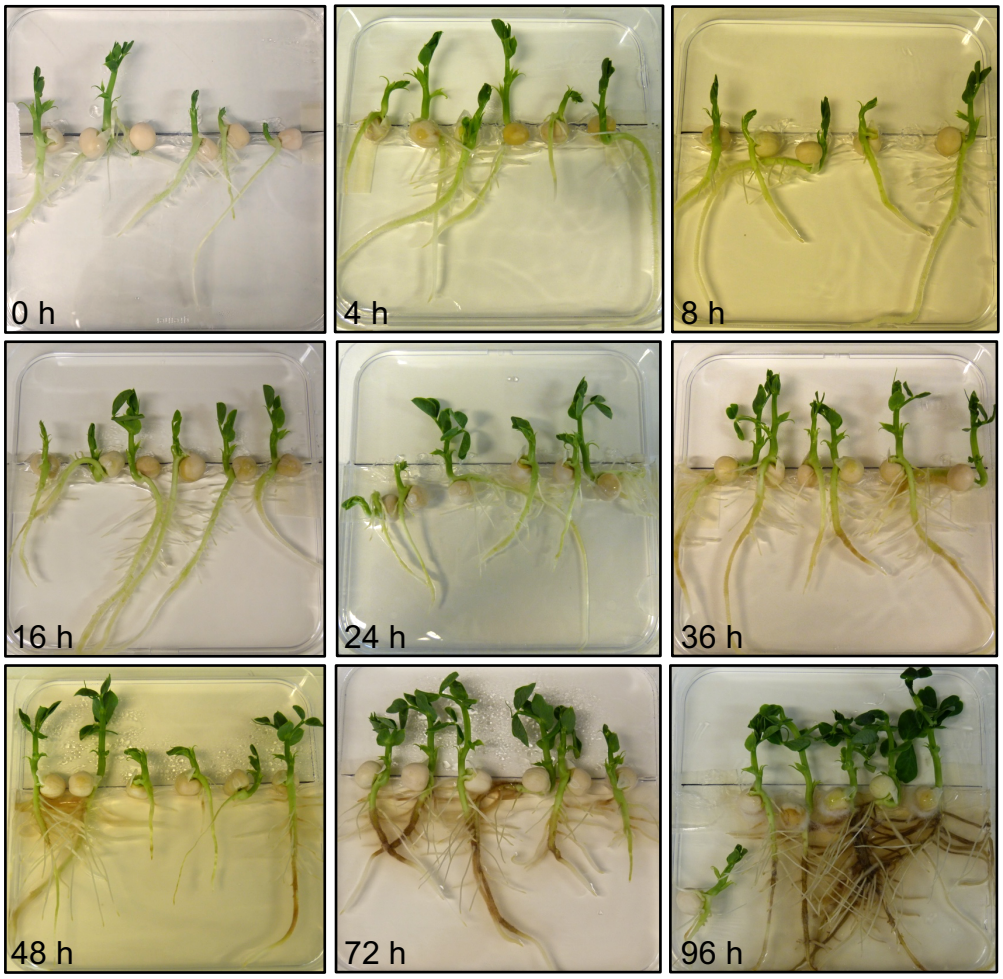


Figure A 4.1 Symptoms of root browning over time on pea seedlings grown in square petri dishes and inoculated with *Fusarium oxysporum* f. sp. *pisi* race 2 isolate F81.

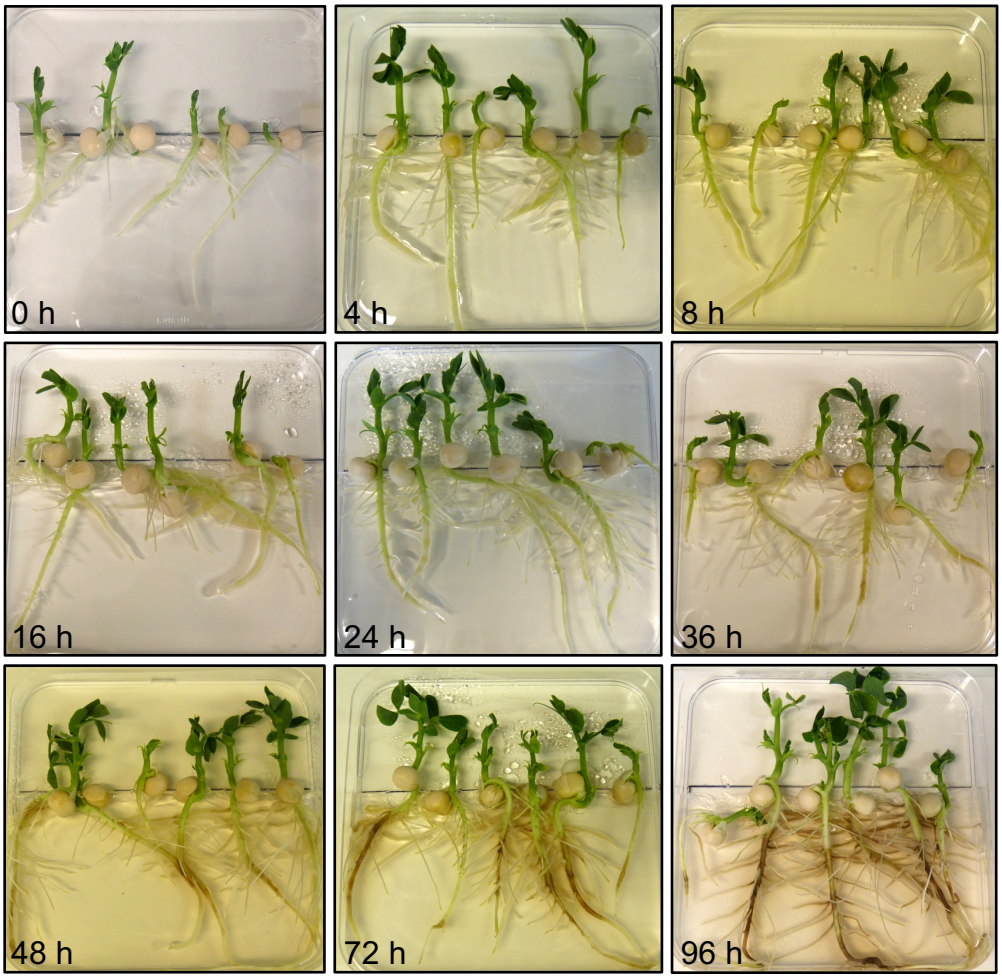


Figure A 4.2 Symptoms of root browning over time on pea seedlings grown in square petri dishes and inoculated with *Fusarium oxysporum* f. sp. *pisi* race 5 isolate R2.

Table A 4.1 Primer pairs used in the characterisation of *Fusarium* species, *SIX* gene presence/absence and *SIX* gene expression studies, with primer name, sequence, annealing temperature and relevant publications included.

Gene	Primer	Sequence 5'-3' (forward/reverse)	Annealing temp. (˚C)	Reference
TEF	exTEF-F/ FUexTEF-R	ACCCGGTTCAAGCATCCGATCTGCGA/ AGCTTGCCRGACTTGATCTCACGCTC	64	Vágány (2012); Taylor *et al.* (2016)
RPB2	7cF/11aR	ATGGGYAARCAAGCYATGGG/GCRTGGATCTTRTCRTCSACC	60	O'Donnell *et al.* (2007)
TUB2	T1/T22	AACATGCGTGAGATTGTAAGT/TCTGGATGTTGTTGGGAATCC	60	O'Donnell *et al.* (1998)
SIX1	SIX1 ALL F1/R1	CTCGGCACCCTCTCAATC/CATTGGTGACAGCATCGTTG	55	This study
SIX6[1]	SIX6 ALL F/R2	TATGCCTGAGCACACCATCAAT/CTCCCAGAGCCATGTATA	50	This study
SIX6[2]	SIX6 FOP1 C481 F/R	AGCACATCAACGAGATCACG/TTTGAAACACCAGGTATAGG	53	This study
SIX7	SIX7 FOL F/R	CATCTTTTCGCCGACTTGGT/CTTAGCACCCTTGAGTAACT	59	Lievens *et al.* (2009)
SIX9	qSIX9 F/R	GCCGACCCAGACCTACGCTTT/GCTGGTTTTGGAAGCCCAGTTGT	63	Dr. A Taylor, unpublished
SIX10	FOC SIX10 F/R	GTTAGCAACTGCGAGACACTAGAA/AGCAACTTCCTTCCTCTTACTAGC	63	Taylor *et al.* (2016)
SIX11	SIX11 FOP1 F/R	CGCAGAGGTTGACCAATAGGTC/CCCAACTTGTTCTGGGGGATTT	61	This study
SIX12	SIX12 FOP1 F/R	GCCGTTCGCACTCCTAGTCATT/CGCATCTCTTCCTTCGCGTACT	63	This study
SIX13	SIX13 FOP2/5 F/R	TCGAAATCCTTCATCATCGACAA/TGTAGCGTTCAAACCACCCTTG	61	This study
SIX14	SIX14 FOL F/R	ATAAAGTGCGACTGGACTTCTGCC/ACCCCCATCCACATTCCTAAGCGA	67	Taylor *et al.* (2016)
Primers used for real time reverse transcription-polymerase chain reaction (qPCR)				
TEF	qTEF F2/R2	GGTCAGGTCGGTGCTGGTTACG/TGGATCTCGGCGAACTTGCAGG	63	Taylor *et al.* (2016)
TUB	qTUB F/R	TTCTGCTGTCATGTCCGGTGT/TCAGAGGAGCAAAGCCAACCA	63	Taylor *et al.* (2016)
SIX1	qSIX1 F3/R2	CGGTCTGTGCGTTGAAAGGTTTG/CTCATTCTCCCCTCGGACATAG	63	This study
SIX6	qSIX6 ALL F3/R2	GAAGGCGAATTTATCATCTTTGG/CTTTTCCCGGTTGCTGCGAG	62	This study
SIX7	qSIX7 F3/R3	TCGATCTCTTTCCAAGACAAGGGCA/GTGGACGCGGCGTTGGTGAAC	63	Taylor *et al.* (2016)
SIX9	qSIX9 F/R	GCCGACCCAGACCTACGCTTT/GCTGGTTTTGGAAGCCCAGTTGT	63	Dr. A Taylor, unpublished
SIX10	qSIX10 F2/R2	CCCGGAAAGCCTGCATCGACTA/AGAACAAACGTCGGTGGGACCA	63	Taylor *et al.* (2016)
SIX11	qSIX11 A F/R	GGCCACACCTGCACGAAAG/CGCAGTTCTTCCCGTCTTTG	60	This study
SIX12	qSIX12 F3/R3	TGCTGCTCCAAGTACAAACTACCTT/GCTGATACCTTTGGGTCCAACGC	63	Taylor *et al.* (2016)
SIX13	FON qSIX13 F2/R2	ACAGCACGGGACAGCTTACA/CGTGAGAGGGGTAGCCACAT	60	Dr. A Taylor, unpublished
SIX14	qSIX14 E F/R	GCTCTGTCTCAGCGTATCCTC/CGACCTGAAACTACCGCCTG	62	This study, designed by Dr. A Taylor

Table A 4.2 The 30 most highly expressed downregulated genes following RNAseq analysis of pea roots infected with *Fusarium oxysporum* f. sp. *pisi* (FOP) isolate FOP1 EMR (race 1) at 96 hpi compared to *in vitro* grown mycelium. Expression levels are ranked in descending order using the average FPKM (fragments per kilobase million) values from three replicates for the 96 hpi data. Genome annotations are details from interproscan which suggest possible associated protein signatures, along with any predicted carbohydrate active enzymes (CAZyme) identified by the CAZY database. Predicted secreted and effector-like proteins (EffP), and numbers of proteins identified as orthologous are included. Orthogroup counts are between the isolates F81(race 2)**:**R2(race 5)**:**F1EMR(race 1)**:**FOC**:**FOL.

Gene name	Contig	Secreted	EffP	Orthogroup contents	Gene annotations	FPKM (96 h)	FPKM (myc.)
g12178	9		Yes			21130	52687
g9787	6		Yes	3:1:2:1:1	protein family UPF0057	6113	16858
g5024	3					2835	6652
g8340	5		Yes	1:1:1:0:0		2315	7921
g9644	6			1:1:1:1:0	Immunoglobulin E-set	1912	9504
g21272	48					1754	5490
g13725	11		Yes			1271	8021
g4065	2		Yes	1:0:1:1:0	Cysteine-rich TM module stress tolerance	1227	5927
g9498	6		Yes	2:1:1:1:2	SCP-2 sterol transfer family	1134	4865
g11777	8		Yes	1:1:1:1:1		1132	3723
g1792	1	Yes	Yes	1:1:1:1:1		1080	8177
g16957	18		Yes	49:16:174:63:2	Chromo (Chromatin Organisation Modifier) domain	987	3583
g14140	11			1:1:1:1:2		885	6654
g7610	5			35:35:34:35:61	Sugar transporter	793	1727
g19817	33		Yes			732	1863
g12065	8		Yes			727	5737
g14363	12			0:1:1:0:0		680	2126
g21300	48			4:2:6:2:6		646	1850
g2752	2			1:1:1:1:1		595	1875
g13726	11		Yes	1:1:1:1:1		594	3387
g14190	11	Yes	Yes	1:1:1:1:1	RlpA-like protein	577	3886
g20555	39			4:2:6:2:6		530	1906
g6071	4			1:2:2:2:2		511	1204
g490	1			3:1:2:2:3	Membrane protein	495	3119
g9358	6		Yes	1:1:1:1:2	RNA recognition motif domain	480	1355
g21792	61		Yes	4:3:3:3:12		471	1809
g8591	5	Yes		1:1:1:1:1		453	1754
g19804	32			1:1:2:0:0		450	1315
g14878	12	Yes	Yes	1:1:1:1:1	protein family UPF0057	437	2020
g2259	1		Yes	0:0:1:1:0		411	867

Table A 4.3 The 30 most highly expressed downregulated genes following RNAseq analysis of pea roots infected with *Fusarium oxysporum* f. sp. *pisi* (FOP) isolate F81 (race 2) at 96 hpi compared to *in vitro* grown mycelium. Expression levels are ranked in descending order using the average FPKM (fragments per kilobase million) values from three replicates for the 96 hpi data. Genome annotations are details from interproscan which suggest possible associated protein signatures, along with any predicted carbohydrate active enzymes (CAZyme) identified by the CAZY database. Predicted secreted and effector-like proteins (EffP), and numbers of proteins identified as orthologous are included. Orthogroup counts are between the isolates F81(race 2)**:**R2(race 5)**:**F1EMR(race 1)**:**FOC**:**FOL.

Gene name	Contig	Secreted	EffP	Orthogroup contents	Gene annotations	FPKM (96 h)	FPKM (myc.)
g12949	12	Yes	Yes	1:1:1:1:1		9699	20081
g6987	6		Yes	3:1:2:1:1	Proteolipid membrane potential modulator	4632	11591
g488	1			1:1:1:1:1		2449	5225
g10154	9			1:1:1:1:1		2031	12506
g371	1		Yes	1:1:1:1:1		1211	6500
g10067	9	Yes		7:6:10:4:10	Cysteine-rich secretory protein family	1162	2964
g172	1	Yes	Yes	1:1:1:1:1		1073	5151
g13814	14					1050	4950
g1880	2		Yes	1:1:1:1:2	Glucose-repressible protein Grg1	925	2469
g14970	16					879	1902
g5008	4			1:1:1:1:1		728	4547
g9324	8	Yes		1:1:1:1:1	Tyrosinase copper-binding domain	669	1407
g6604	6			2:2:1:1:6	Hyaluronan/mRNA-binding protein	599	1235
g6604	6			2:2:1:1:6	Hyaluronan/mRNA-binding protein	599	1235
g14418	15	Yes	Yes	1:1:1:0:0		596	1548
g9800	9	Yes	Yes	1:1:1:1:1		552	1222
g15444	16			1:1:1:1:1		525	1203
g4035	3		Yes			524	7244
g202	1			1:1:1:1:1		481	1262
g5145	4	Yes		1:1:1:1:2	Growth factor receptor cysteine-rich domain	453	1658
g7024	6			1:1:1:1:1	Unknown function	450	1339
g555	1			2:1:1:1:7	Pleckstrin homology domain	438	923
g711	1				Membrane protein, TMD	409	1103
g4665	4			1:0:0:0:1		402	981
g7205	6		Yes	1:1:1:1:0		384	793
g712	1		Yes	2:4:1:1:2		383	1022
g15165	16		Yes	101:72:106:52:16		381	1328
g16158	18			1:1:1:1:1	Phospholipase, TMD	367	1067
g9422	8			1:1:1:1:1	Membrane protein	356	1321
g4773	4		Yes	1:1:1:1:2		350	3243

Table A 4.4 The 30 most highly expressed downregulated genes following RNAseq analysis of pea roots infected with *Fusarium oxysporum* f. sp. *pisi* (FOP) isolate R2 (race 5) at 96 hpi compared to *in vitro* grown mycelium. Expression levels are ranked in descending order using the average FPKM (fragments per kilobase million) values from three replicates for the 96 hpi data. Genome annotations are details from interproscan which suggest possible associated protein signatures, along with any predicted carbohydrate active enzymes (CAZyme) identified by the CAZY database. Predicted secreted and effector-like proteins (EffP), and numbers of proteins identified as orthologous are included. Orthogroup counts are between the isolates F81(race 2)**:**R2(race 5)**:**F1EMR(race 1)**:**FOC**:**FOL.

Gene name	Contig	Secreted	EffP	Orthogroup contents	Gene annotations	FPKM (96 h)	FPKM (myc.)
g4853	3		Yes			7005	76085
g5974	4		Yes	3:1:2:1:1	Proteolipid membrane potential	6854	14857
g4824	3		Yes	1:1:1:1:1		4175	27417
g12378	9			1:1:1:1:1	Zinc finger C2H2 type domain profile, TF	4024	9159
g14603	11					2454	13849
g595	1		Yes	1:1:1:1:1		2111	10106
g10932	8			1:1:1:1:1		1614	10734
g4113	3	Yes	Yes	0:1:0:1:2		795	48699
g14921	12			0:1:0:1:0		763	1846
g12092	9		Yes	1:1:1:1:1	Basic-leucine zipper, TF	737	3426
g6402	4		Yes	1:1:1:1:2	RNA recognition motif domain	703	1618
g16707	15		Yes	0:1:0:1:0		666	2190
g422	1			1:1:1:1:1		665	1394
g10800	8			0:1:1:0:0		657	1686
g4213	3	Yes	Yes	1:1:1:1:1		630	1964
g14353	11			1:1:1:1:2	C2H2 and C2HC zinc fingers, TF	558	2343
g4454	3			3:1:2:2:3	fungal rhodopsins	535	1871
g9907	7		Yes	1:1:1:1:2		510	3813
g784	1			2:1:1:1:7	Pleckstrin homology domain	498	1078
g12916	10			4:2:6:2:6	Unknown	487	1101
g12915	10		Yes	1:1:1:1:1		478	1036
g1471	1		Yes	1:1:1:1:1		472	1122
g201	1			1:1:1:1:1	Cardiolipin synthase N-terminal, TMD	467	2385
g14923	12			0:1:0:1:0		433	1017
g4854	3		Yes	1:1:1:1:3		430	1225
g1350	1			1:1:1:1:1	Unknown function, membrane bound	391	1015
g14224	11				Membrane protein, TMD	389	1295
g2626	2			1:1:1:1:1	Membrane protein	381	2225
g2972	2		Yes	1:1:1:1:1		379	906
g4214	3					377	1148