AIX-MARSEILLE UNIVERSITÉ

Ecole Doctorale Sciences de la Vie et de la Santé

THÈSE DE DOCTORAT

Spécialité: Génétique

Presenté par: Le HE

Pour obtenir le grade de docteur de l'Université Aix-Marseille

Interactions hôte-pathogène entre *Caenorhabditis elegans* et le champignon *Drechmeria coniospora*

Soutenue le 2 décembre 2016 devant le jury composé de:

Dr. Dominique Ferrandon	Rapporteur
Dr. Hinrich Schulenburg	Rapporteur
Dr. Philippe Naquet	Président
Dr. Eric Record	Invité
Dr. Jonathan Ewbank	Directeur de thèse

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

CHAPTER 1. INTRODUCTION

1.1 Host-pathogen interactions

1.1.1 *C. elegans* and its innate immunity

C. elegans is a small (around 1 mm in length), free-living worm, generally found in rotting fruit and plant matter. It feeds on bacteria and fungi. In the laboratory, it can be maintained on agar plates with a spread layer of *Escherichia coli* as its food. Starting in the mid-1960s, Sydney Brenner developed *C. elegans* as a model organism, for studying developmental biology and neurobiology (BRENNER 1974). It takes 3 days for the worm to develop from embryo to adult at 20 °C. During its adulthood, a single hermaphrodite can produce around 300 genetically-identical progeny. The hermaphrodite can also be fertilized by males. Males arise at a very low frequency, 0.1% under standard laboratory conditions. The life span of *C. elegans* is about 2-3 weeks on average at 20 °C. *C. elegans* was the first multicellular organism to be sequenced (THE C. ELEGANS SEQUENCING CONSORTIUM 1998). Strains are generally of the same, defined, genetic background. Loss-of function mutants for most genes are available. Additionally, one can knock-down most of its genes with RNAi by feeding (further introduced in 1.2.1.1). All these factors have made *C. elegans* a good model for various biological studies. Pioneered by Ausubel's lab, *C. elegans* has also emerged as a model organism for the study of host-pathogen interactions in the last 15 years (MAHAJAN-MIKLOS *et al.* 1999; TAN *et al.* 1999).

There are 3 layers for *C. elegans* defence against pathogens. Firstly, worms are capable of distinguishing some pathogens from non-pathogenic microbes and move away

from them, thus potentially avoiding infection (SCHULENBURG AND EWBANK 2007; MCMULLAN *et al.* 2012; NAKAD *et al.* 2016). In addition, there are two physical barriers that protect worms from pathogenic microbes. A tough cuticle of collagen covers the whole worm's body and keeps microbes away from its epidermis. The grinder within the pharynx is another barrier that destroys most microbes before they enter the worm intestine (KIM AND EWBANK 2015). Last but not least, the worm has an immune system. Lacking adaptive immunity, as for other invertebrates, it relies exclusively on its innate immunity to combat infection. Some aspects of this innate immune system are conserved and are present in a broad range of species, from plants to vertebrates (KURZ AND EWBANK 2003; MILLET AND EWBANK 2004; NICHOLAS AND HODGKIN 2004; SCHULENBURG *et al.* 2004; IRAZOQUI *et al.* 2010). To combat an infection efficiently, the worm needs to detect the microbe, generally through the recognition of conserved features, or needs to sense the damage caused by an invading pathogen, so-called microbe associated molecular patterns (MAMPs) or damage associated molecular patterns (DAMPs), respectively. Both MAMPs and DAMPs are recognized by pattern recognition receptors (PRRs) that activate immune signalling pathways leading to defence gene expression.

Several families of potential antimicrobial effectors proteins have been identified through the investigation of the changes in host gene expression that accompany infection in worms. These studies have used bacterial pathogens, including the Gram-negative bacterium *Serratia marsescens* (MALLO *et al.* 2002), and the obligate nematode fungal pathogen *Drechmeria coniospora* (COUILLAULT *et al.* 2004; PUJOL *et al.* 2008). They

Table 1.1 NLP and CNC genes and their induction by different stimuli

Gene	Constitutive expression	Relative induction		
		D. coniospora infection	Wounding	Osmotic stress
nlp-27	+++	+	+	-
nlp-28	+	+++	+++	++
nlp-29	+	+++	+++	++
nlp-30	+	++	++	-
nlp-31	-	+++	+++	-
nlp-34	-	+++	+++	-
cnc-1	-	+++	+++	+
cnc-2	+	+++	+	+
cnc-3	+	-	-	+
cnc-4	+	+++	++	+
cnc-5	-	+++	+++	+
cnc-11	-	+	+++	++

Data are based on (PUJOL *et al.* 2008; ZUGASTI AND EWBANK 2009). The expression of the *nlp* (*cnc*) genes was compared with that of *nlp-29* (*cnc-2*). For constitutive expression, "-" means gene not expressed; for relative induction, "-" means no expression change.

include saposin-like proteins (SPPs), neuropeptide-like proteins (NLPs) and Caenacins (CNCs). Several of the NLPs that are antimicrobial peptides (AMPs) are encoded by a gene cluster (including *nlp-27*, *nlp-28*, *nlp-29*, *nlp-30*, *nlp-31* and *nlp-34*) and share a YGGYG sequence motif. CNCs are another group of AMPs that are structurally similar to NLPs. The best studied are also found as a cluster in the genome, *cnc-1*, *cnc-2*, *cnc-3*, *cnc-4*, *cnc-5*, and *cnc-11* (COUILLAULT *et al.* 2004; PUJOL *et al.* 2008; PUJOL *et al.* 2012). These two families of AMPs contribute to the resistance of *C. elegans* to *D. coniospora* infection and some of them are also differentially expressed upon other stresses such as injury and high osmolarity (Table 1.1). These differences in expression upon different stresses reflect the underlying complex regulatory network (EWBANK 2006; PUJOL *et al.* 2012; ZUGASTI *et al.* 2016). Although some genes are required for the regulation of AMP expression following both osmotic stress and infection, the two responses are genetically distinct (LEE *et al.* 2010; ZUGASTI *et al.* 2016). Among the elements required for the response to infection, DCAR-1 is a DAMP receptor that was recently discovered through a reverse genetic screen. This receptor is expressed in epidermis where it can sense 2-hydroxy-3-(4-hydroxyphenyl) propanoic acid (HPLA), a DAMP generated by *C. elegans* after *D. coniospora* infection or cuticle damage. It regulates AMPs via a conserved p38 mitogen-activated protein kinase (MAPK) pathway (ZUGASTI *et al.* 2014). Currently, *dcar-1* has only been shown to be important for the resistance of *C. elegans* to *D. coniospora*, but it is likely to be involved in defence against other nematophagous fungi.

Table 1.2 Typical infection structures of some nematophagous fungi (adapted from NORDBRING-HERTZ *et al.* 2011)

Infection structure	Species	Taxonomic classification	Sequenced
Adhesive nets	*Arthrobotrys oligospora*	Ascomycota; Orbiliales	Yes
	A. conoides		
	A. musiformis		
	A. superba		
	Duddingtonia flagrans		
Adhesive branches	*Monacrosporium gephyropagum*	Ascomycota; Orbiliales	
Adhesive knobs	*M. ellipsosporum*	Ascomycota; Orbiliales	
	M. haptotylum		Yes
Constricting rings	*A. dactyloides*	Ascomycota; Orbiliales	
	A. brochopaga		
Adhesive knobs and adhesive spores	*Nematoctonus concurrens*	Basidiomycota; Agaricales	
Adhesive spores	*N. leiosporus*		
	Drechmeria coniospora	Ascomycota; Hypocreales	Yes
	Hirsutella rhossiliensis	Ascomycota; Hypocreales	
	H. minnesotensis		Yes
Ingested spores	*Harposporium anguillulae*	Ascomycota; Hypocreales	
Zoospores	*Catenaria anguillulae*	Chytridiomycota; Blastocladiales	
Adhesive hyphae	*Haptoglossa dickii*	Oomycota; Haptoglossales	
	Stylopage hadra	Zygomycota; Zoopagales	
	Cystopage cladospora		
Toxic droplets	*Pleurotus ostreatus*	Basidiomycota; Agaricales	
Appressoria	*Pochonia chlamydosporia*	Ascomycota; Hypocreales	Yes
	Paecilomyces lilacinus	Ascomycota; Eurotiales	

1.1.2 Nematophagous fungus

Nematophagous fungi are nature enemies of nematodes. More than 200 species have been described that all are capable of attacking living nematodes. In general, these fungi can be grouped into three categories according to their strategies for approaching the nematodes. Firstly, trap-forming fungi can produce sophisticated mycelial nets, rings, knobs or branches by which they mechanically capture the nematode or trap them with specialised adhesive structures. Secondly, endoparasites rely on spores that can infect nematodes through the cuticle or intestine. Lastly, there are fungi that use their vegetative hyphae to infect eggs or the sedentary females of root-knot nematodes (Table 1.2).

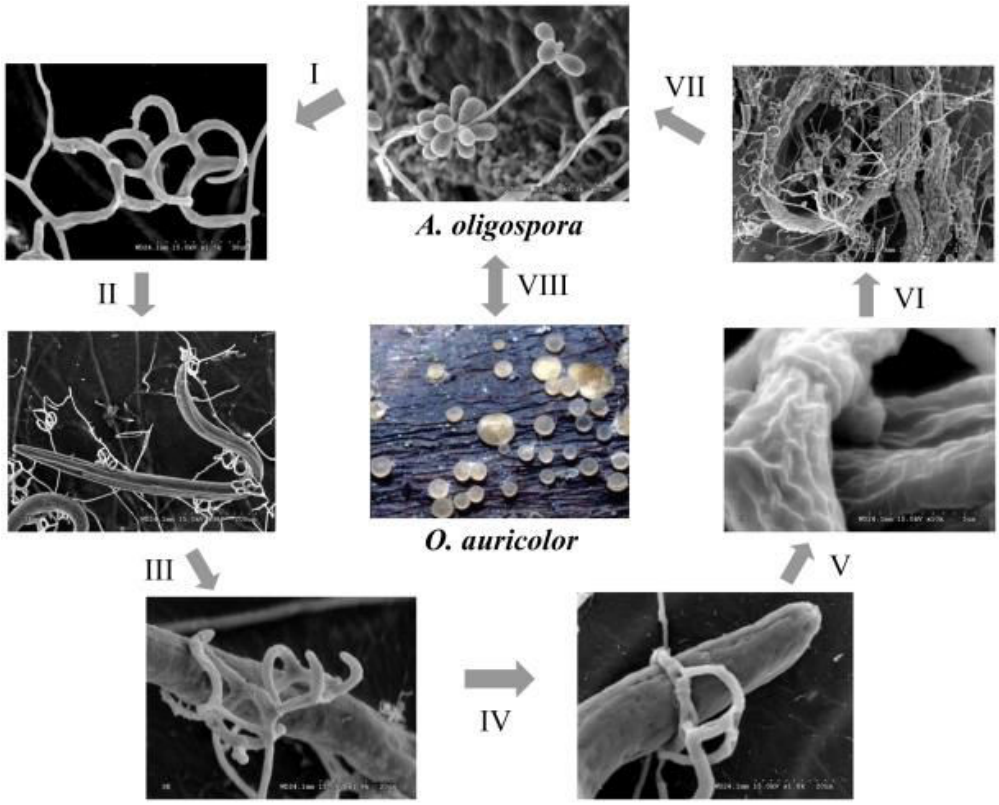


Figure 1.1 Saprophytic and parasitic stages of the nematode-trapping fungus Arthrobotrys oligospora. (copied from Yang et al. 2011)

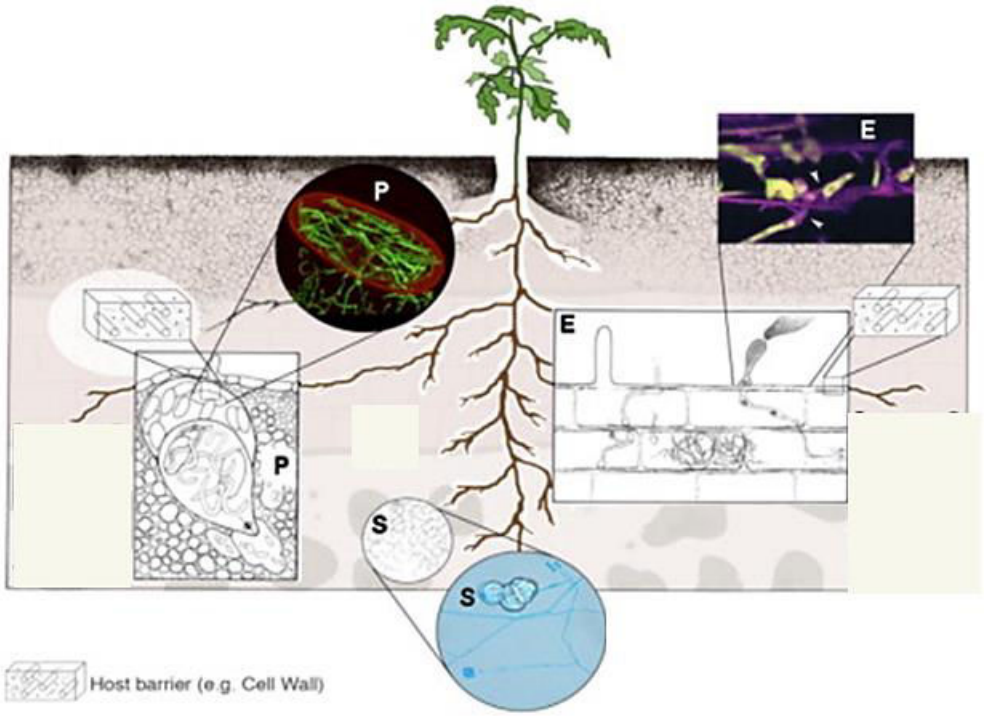


Figure 1.2 Tritrophic lifestyles of the nematophagous fungus *P. chlamydosporia* (copied from LARRIBA *et al.* 2014)

Arthrobotrys spp. such as *A. oligospora* and *A. conoides* belong to the first group. Fungi in this group possess the ability to survive in soil saprophytically. Among these fungi, *A. oligospora* is one of the best-studied nematode-trapping fungi and it was also the first to be sequenced (YANG *et al.* 2011). This fungus lives as a saprophyte in diverse soil environments where it mainly lives. In the presence of nematodes, it undergoes morphological change and forms 3D traps to capture its prey (Figure 1.1 I, II). The adhesive traps subsequently penetrate and immobilize the worm (Figure 1.1 III, IV, V). Finally, the fungus digests the worm, and then goes onto another round of infection, reverts to its saprophytic lifestyle (Figure 1.1, VII) or enters a sexual stage; its teleomorph is named as *Orbilia auricolor* (Figure 1.1, VIII). Similarly, another sequenced fungus *Monacrosporium haptotylum* traps and infects the nematode with its adhesive knob (AHREN *et al.* 2005; MEERUPATI *et al.* 2013).

P. chlamydosporia infects nematode eggs with its infective appressorium, formed after the direct contact with a worm egg. It can bind to the eggshell and penetrate inside the egg to mount an infection (OLIVARES AND LOPEZ-LLORCA 2002). Besides the pathogenic life in worms (P, indicated in Figure 1.2), this fungus can live in two other ways, as a saprophyte (S, Figure 1.2) and as a root endophyte of plant (E, LARRIBA *et al.* 2014).

The obligate nematode pathogen *D. coniospora* is an endoparasite. It has been used to study *C. elegans* innate immunity to fungal infection for more than 10 years (COUILLAULT *et al.* 2004). This fungus was first discovered by Drechsler (DRECHSLER 1941) and its infection process was described by Jansson (JANSSON AND NORDBRING-

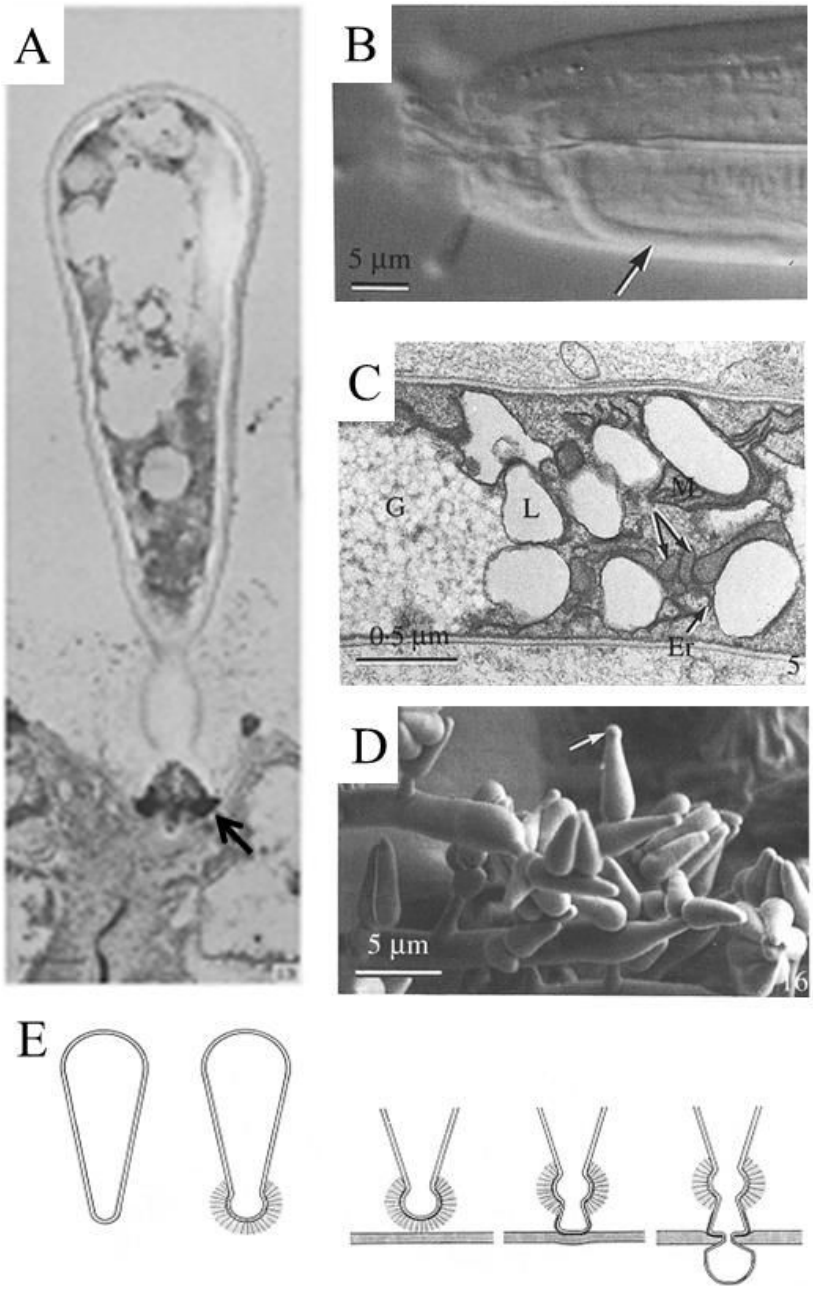


Figure 1.3 Parasitic lifecycle of *D. coniospora*. (A). Electron micrograph on cytochemical stained infection site showed strong acid phosphatase activity at the penetration site (arrow). (B). Light microscope shows the hypha inside worm after penetration. (C). Electron microscopic image shows lipid droplets L, glycogen G, microbodies M (double heads arrow) and endoplasmic reticulum Er (arrow) inside fungus. (D). Spore clusters from infected dead worm. The arrow is pointing at an adhesive bud. E. Schematic representation of the initial stage of infection. Figures were adapted from DIJKSTERHUIS *et al.* 1990 and DIJKSTERHUIS *et al.* 1991

HERTZ 1983; JANSSON *et al.* 1984; JANSSON 1994) and Dijksterhuis (DIJKSTERHUIS *et al.* 1990; DIJKSTERHUIS *et al.* 1991). For the initial steps of infection, non-motile spores (also called conidia) released from spore clusters develop an adhesive bud that binds to the worm cuticle. A visible appressorium will subsequently form between the bud and cuticle. With a combination of turgor pressure and biochemical activity (as indicated by histochemical staining of acid phosphatase activity), the germinating hyphae wi'' nake a narrow penetration plug to get into the worm (DIJKSTERHUIS *et al.* 1990; DIJKSTERHUIS *et al.* 1991). The trophic hyphae inside the worm form a large number of lipid droplets and microbodies (Figure 1.3 B,C). *C. elegans* dies some 48 h after initiation of infection. Hyphae grow out from the worm cuticle and generate spore clusters, thus completing the life cycle (Figure 1.3 D). Complete genome sequences are now available for 2 independent isolates of *D. coniospora* (LEBRIGAND *et al.* 2016; ZHANG *et al.* 2016). Another sequenced species *Hirsutella minnesotensis* also infects nematodes with similar infection process, but with different way bind to the host. Without forming an adhesive bud, *H. minnesotensis* adheres to any part of the worm cuticle with its secreted adhesive substances (CHEN *et al.* 2000).

1.1.3 Plant pathogenic fungi

There are parallels between the infection of *C. elegans* by *D. coniospora* and of plants by fungal pathogens. Firstly, the fungal spores bind to cuticle of the worm or plant, then they germinate and form the infecting structure, the appressorium. For plant fungal pathogens, appressorium formation is a response of the spore to the cuticle's chemical composition or physical features (O'CONNELL AND PANSTRUGA 2006). The nematode-trapping fungus

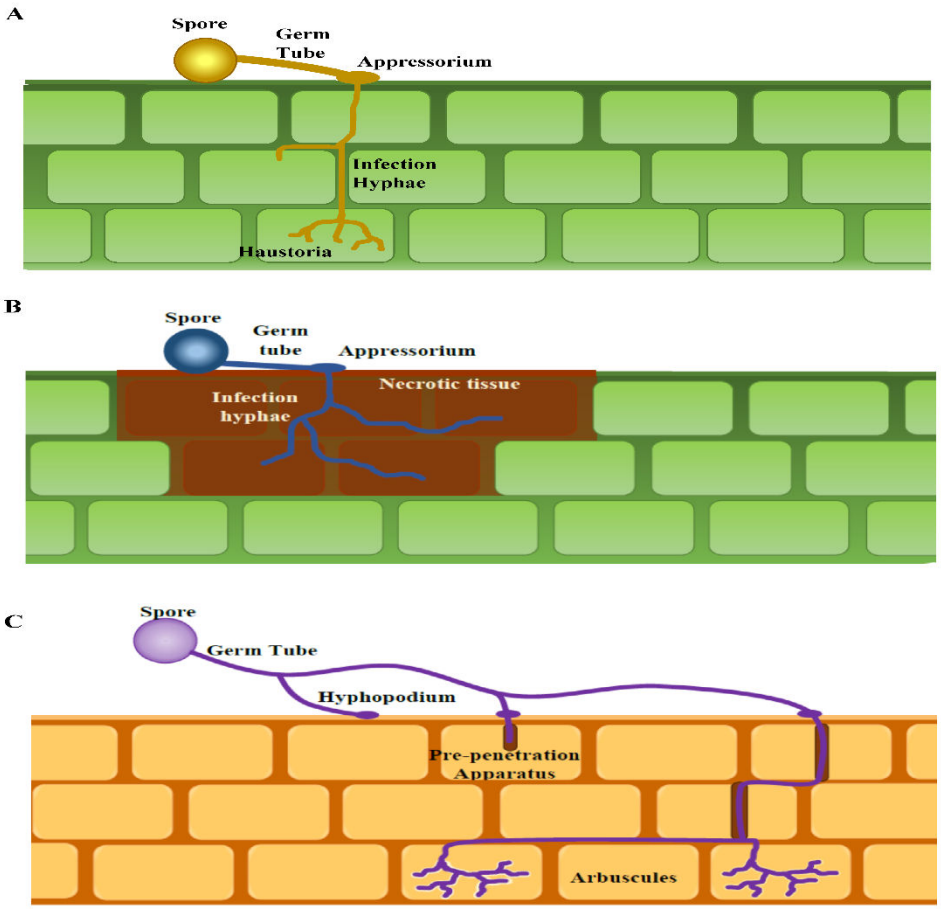


Figure 1.4 Diagram of pathogenic lifestyles. (A). Biotrophic pathogen; (B). Necrotrophic pathogen; (C). Hemitrophic pathogen. Copied from MICHELLE AND HARTMAN 2016.

Arthrobotrys oligospora can detect nematode pheromones and form abundant traps (HSUEH *et al.* 2013). The search for the stimulus triggering *D. coniospora* appressorium formation is, however, still on going.

The infectious hyphae that grow in a host plant from the maturing appressoria secrete cell-wall degrading enzymes (CWDEs) that allow the cell wall to be breached and the hyphae to enter. In general, CWDEs include cellulose-degrading enzymes, hemicellulose-degrading enzymes, pectin-degrading enzymes, side-chain cleaving and other accessory enzymes (KUBICEK *et al.* 2014). Necrotrophs tend to secrete high concentrations of CWDEs, but biotrophs such as *Magnaporthe oryzae* use a combination of CWDEs and turgor pressure from appressoria to penetrate through the cell wall without causing host cell death (TUCKER AND TALBOT 2001). *D. coniospora* forms a visible appressorium and uses the same strategy of combining turgor pressure and enzymatic activity to bore into the worm (DIJKSTERHUIS *et al.* 1990). The thick collagen-containing cuticle, which is functionally similar to the plant cell wall, protects nematodes from many pathogens. Its composition and structure changes during worm development and certain pathogens may only be capable of breaching it at specific developmental stages. *Both H. minnesotensis* and *A. oligospora* are thought to use pepsin-like aspartic peptidases and subtilisin serine peptidases to infect worms through the cuticle (YANG *et al.* 2011; LAI *et al.* 2014). *P. chlamydosporia*, on the other hand, degrades and penetrates into the nematode chitin-rich egg-shell with a battery of secreted hydrolytic enzymes, including proteases, chitinases, esterases and lipases (LARRIBA *et al.* 2014).

Fungal colonization follows the successful penetration through the plant cell wall. Like the penetration tactics, the colonization strategies also differ from fungi according to their lifestyle. The necrotrophic fungi, such as *Botrytis cinerea,* grow under the cuticle and kill the plant epidermal cells with secreted toxic metabolites and proteins. The biotrophs can grow inside or between the plant cells without killing them. Notably, the hemibiotrophic fungi like *M. oryzae* and *Colletotrichum spp*. first invade the host biographically; later on they change to a necrotrophic stage and kill the host cells (Figure 1.4). *D. coniospora* probably infect worms in a hemibiotrophic manner, as *C. elegans* epithelial cells are still alive and capable of mounting a defence response during fungal colonization and start to die after 24 h of infection (COUILLAULT *et al.* 2004).

The repertoire of effectors of each species reflects its interaction with its hosts and determines the fungal lifestyle. Compared to the many effectors from bacteria, relatively few fungal effectors have been functionally validated. This is due in part to the difficulty of genetic transformation for obligatory fungal pathogens. There are, however, exceptions and fungal effectors have been functionally characterized from several plant-pathogen infection models (Table 1.3; LO PRESTI *et al.* 2015).

Table 1.3 Example of selected functionally verified fungal effectors in plant pathogenic fungi

Fungal species, lifestyle	Effectors, localization	Host counterpart, function for defence	reference
Ustilago maydis, biotroph	Pep1, apoplast	POX12, a secreted maize peroxidase, a component of the plant reactive oxygen species (ROS)- generating system	(HEMETSBERGER *et al.* 2012)
	Pit2, apoplast	Cysteine proteases, promotes salicylic acid associated plant defence	(MUELLER *et al.* 2013)
	Cmu1 (chorismate mutase, reduces the level of chorismate), cytosol	Chorismate, precursor for the synthesis of salicylic acid, a necessary component for the salicylic acid associated plant defence	(DJAMEI *et al.* 2011)
	Tin2, cytosol	ZmTTK1, boost the plant cell wall lignification and impose a physical barrier for pathogen spread.	(TANAKA *et al.* 2014)
Cladosporium fulvum, biotroph	Avr4 (binds to fungal chitin on its cell wall), apoplast	Chitinase, hydrolyses the fungal cell wall	(VAN DEN BURG *et al.* 2006)
	Ecp6 (sequesters released chitin oligosaccharides), apoplast	Pathogenic chitin-induced immune reaction	(SANCHEZ-VALLET *et al.* 2013)
	Avr2, apoplast	PIP1/Rcr3, apoplastic proteases	(ROONEY *et al.* 2005; SHABAB *et al.* 2008)
	Tom1(tomatinase, degrades α-tomatine into β-tomatine and tomatidine), apoplast	α -tomatine, Antifungal glycoalkaloid	(OKMEN *et al.* 2013)

Fungal species, lifestyle	Effectors, localization	Host counterpart, function for defence	reference
Blumeria graminis f. sp. *hordei,* biotroph	CSEP0005, apoplast	Plant pathogenesis-related 1 (PR1) and PR 17 protein family	(ZHANG *et al.* 2012)
	Blumeria effector candidate 4 (BEC4, inactivate ARF-GAP), apoplast	ADP ribosylation factor-GTPase-activating protein (ARF-GAP), defence-related host vesicle trafficking	(SCHMIDT *et al.* 2014)
Stagonospora nodorum and *Pyrenophora triticirepentis*, necrotroph	ToxA (host specific toxin, HST, inactivate ToxABP1), chloroplast	ToxABP1, Thylakoid formation, ToxABP1 draining out leads to chloroplast perturbation and cell death	(MANNING *et al.* 2007)
Botrytis cinerea, necrotroph	Small RNAs	AGO1, immune genes (Introduced in 1.2.1.3)	(WEIBERG *et al.* 2013).
Magnaporthe oryzae, hembioitroph	Slp1 (closely related to Ecp6 in *C. fulvum*), apoplast	CEBiP, PRR chitin elicitor-binding protein, together with OsCERK is important for eliciting pathogenic chitin-induced immune reaction	(MENTLAK *et al.* 2012)
	AvrPiz-t (its R gene in host is Piz-t), cytosol	APIP6 RING E3 ubiquitin ligase, activate ROS generation	(PARK *et al.* 2012)
Colletotrichum higginsianum, hemibiotroph	ChNLP1 (necrosis- and ethylene-inducing peptide 1 like protein, causing host cell death), cytosol		(YOSHINO *et al.* 2012)

A comparison of effectors from fungal pathogen with different lifestyles reveals a range of virulence strategies. Among the biotrophs, effectors can suppress host immune reactions directly. For example, in *Ustilago maydis*, Pep1 is sent to the apoplast to block the activity of host POX12, a component of a ROS-generating defence system (HEMETSBERGER *et al.* 2012); Tom1 in *Cladosporium fulvum* blocks the antifungal agent α-tomatine (OKMEN *et al.* 2013). Effectors also can manipulate host immunity in an indirect way, such as *U. maydis* Cmu1, a chorismate mutase secreted into the host cytosol. It can reduce the amount of chorismate within host cells resulting in a decrease of salicylic acid, and thereby the down regulation of salicylic acid-related immunity (DJAMEI *et al.* 2011). Pit2, another *U. maydis* effector, inhibits the same immune pathway as Cmu1, but in this case acts by functioning as a cysteine proteases inhibitor to suppress the salicylic-acid-associated plant defences.

Fungi can also shield themselves from host PRRs to limit the immune response. Thus, normally, host immunity can be elicited by chitin oligomers via PRRs (MACHO AND ZIPFEL 2014). To avoid this host defence, the biotrophic fungus *C. fulvum* applies an effector-Ecp6 to sequester its own chitin and avoid PRR-based detection (SANCHEZ-VALLET *et al.* 2013). Similarly, Avr4 prevents fungal chitin from being released into the apoplast – the area surrounding the plasma membrane (VAN DEN BURG *et al.* 2006). There are many other mechanisms used by pathogenic fungi to down regulate the host immune reaction, as exemplified by Tin2, Avr2, CSEP0005 and BEC4 that interact with different host immunity-related proteins (Table 1.3).

Necrotrophic effectors that induce host cell death are often toxic metabolites, or toxins, such as necrosis-and-ethylene-inducing peptide 1 (Nep1) or Nep1-like proteins (QUTOB *et al.*

2006). ToxA* is a host-specific toxin that interacts with ToxABP1 in chloroplasts. It causes the disorganization of both photosystem I and II and finally cell death (MANNING *et al.* 2007). Another interesting confirmed necrotrophic effector is a trans-species sRNA, which is described below (section 1.2.1.3). The hemibiotrophic fungi need first to suppress host immunity when growing as biotrophs, but switch to a necrotrophic lifestyle and kill the host cell later. In *Magnaporthe oryzae*, Slp1 and AvrPiz-t both are highly expressed in the early stages of infection, to avoid eliciting chitin-related immunity or a ROS reaction (MENTLAK *et al.* 2012; PARK *et al.* 2012). In *Colletotrichum higginsianum*, ChNLP1- a Nep1-like protein is expressed when the fungus switches to a necrotrophic stage. This protein induces cell death in *Nicotiana benthamiana* (YOSHINO *et al.* 2012). Despite the extensive studies that have been conducted on the interactions between *C. elegans* and fungal pathogens, and the availability of annotated genomes for several nematophagous fungi, there are still a very limited number of effectors that have been reported to be important for pathogenesis.

1.2 Fungal genetic modification

1.2.1 Small RNA for cross-species gene silencing

1.2.1.1 RNA interference and RNAi by feeding in *C. elegans*

One reason for the lack of functional dissection of fungal virulence has been the lack of tools for knocking-down gene function. The technique of RNA interference (RNAi) has several advantages, including the fact that it does not require any modification of the target organism's genome. In RNAi, double-stranded RNA (dsRNA) inhibits the expression of a specific gene on the transcriptional or translational level. In 1998, Andrew Fire and Craig Mello discovered that

* ToxA from *Stagonospora nodorum* and *Pyrenophora triticirepentis* is structurally and functionally unrelated to the well-known bacterial effector of the same name

injecting the dsRNA into worms can trigger the degradation of specific mRNAs (FIRE *et al.* 1998). Other than injection, there are two other ways for introducing double-stranded RNA (dsRNA) into *C. elegans* and triggering RNAi, RNAi by feeding (RNAi feeding) or dsRNA soaking. If worms are fed with bacteria expressing dsRNA, the dsRNA will be absorbed by the worms' intestinal cells after ingestion and spreads to other tissue causing the gene silencing locally (FIRE *et al.* 1998; TIMMONS AND FIRE 1998). Because of its convenience and low cost, RNAi feeding is a widely-used reverse genetic tool for studying gene function, including through genome-wide screening.

To explain how dsRNA triggers RNAi, here I take RNAi feeding as an example. The underlying mechanism is similar however the exogenous dsRNA is introduced. Once the exogenous dsRNA is released from bacteria and gets into the worm intestinal cells, it will spread into the different non-neuronal tissues via the proteins SID-1 and SID-2. A primary exogenous small interfering RNA (siRNA) is generated from the dsRNA by the action of the conserved RNase III enzyme Dicer (DCR-1 in *C. elegans*; ZAMORE *et al.* 2000; BERNSTEIN *et al.* 2001). Afterwards, these primary siRNAs will be loaded onto the Argonaut protein RDE-1 and subsequently bind to their mRNA target(s) inhibiting translation. It is should be noted that the primary siRNA production was also RDE-1 dependent (JOSE *et al.* 2011). In addition to binding on the mRNA, the primary siRNA-RDE-1 complex also triggers the production of secondary siRNA (22G) by recruiting RNA dependent RNA polymerases (RdRP) to amplify the secondary siRNA to propagate the RNAi effect (SIJEN *et al.* 2001). Apart from translational silencing, RNAi also acts at the transcriptional level, guided by the 2^{ry} siRNA. The newly synthesized 2^{ry} siRNA can bind to NRDE-3 or to a WAGO Argonaut protein and enter the nucleus. This complex will associate with the targeted nascent pre-mRNA and further recruit other proteins to

inhibit Pol II elongation and deposit the repressive chromatin mark H3K9me3 (BILLI *et al.* 2014). The epigenetic marks will decrease the accessibility to DNA for RNA polymerase, silencing the targeted gene at transcriptional level.

1.2.1.2 RNAi in fungus

RNAi in fungus was firstly described as quelling in *Neurospora crassa* (ROMANO AND MACINO 1992), 6 years before the RNAi mechanism was clarified in nematode (FIRE *et al.* 1998). As a conserved biological process, the fungal RNAi pathway is similar to that within worms (see section 1.2.1.1). Thus, there are also 3 essential proteins for the fungal RNAi pathway: the RNase III enzyme Dicer is needed for processing long dsRNA or endogenous RNA into primary siRNA; Argonaut proteins will bind to 22-25 nucleotide siRNA and be guided to the targeted mRNA; RdRP is necessary for propagating the RNAi signal.

Although RNAi is conserved across different domains of life, interestingly, this pathway has been lost in some fungi, such as *Saccharomyces cerevisiae* (budding yeast). *S. cerevisiae* lacks homologues of several RNAi components and has an impaired RNAi pathway (ARAVIND *et al.* 2000)**.** This extreme case may be caused by a killer virus. Yeast hosting this virus will produce a toxin to kill the nearby cells that do not contain the virus. If a yeast cell has strong RNAi activity, it will produce siRNA against the virus and therefore not be infected, but will be susceptible to viral toxin produced by neighbouring cells that harbour the virus (DRINNENBERG *et al.* 2011).

1.2.1.3 Cross species RNA interference

The spreading of RNAi between different cells and tissues in one organism relies on the mobility of the small RNAs. Studying with plants shows the mobile RNAs are mainly duplex 21-

nucleotide-long (21-nt-long) and 24-nt-long siRNA. These duplex siRNAs are transported from the original tissue through plasmodesmata and phloem to the other parts of the plant and transmit the gene silencing to these parts. These two different small RNA (sRNA) species spread the silencing effect broadly inside the plant (DUNOYER *et al.* 2010; MOLNAR *et al.* 2010).

C. elegans uptakes dsRNA efficiently from the environment, as showed by RNAi feeding. The mobile RNA in worms from RNAi feeding can be divided into 2 types: the ingested exogenous long dsRNA and small RNAs generated by the worm. As mentioned before, the SID-1 and SID-2 protein are responsible for transferring the dsRNA, SID-1 particularly is preferred for transferring long dsRNA (WINSTON *et al.* 2002). These long dsRNAs are transferred between worm tissues and cleaved by DCR-1 in an RDE-1 independent manner; the DCR-1 processed intermediate primary siRNA is mobile (JOSE *et al.* 2011).

Based on the abundance of RNAs between the host and pathogen during infection, RNA-based communication could be another aspect of host-pathogen interactions. Indeed, several cases of cross-species RNAi have been described, involving sRNA trafficking between a host and its pathogen. *Botrytis cinerea*, a fungal plant pathogen, selectively secretes specific siRNAs into *Arabidopsis* and silences host defence genes: mitogen-activated protein kinase genes MPK1 and MPK2, a cell wall-associated kinase (WAK), a peroxiredoxin (PEXIIF) and the tomato MPK-kinase kinase 4 (MAPKKK4). The host Argonaut 1 is necessary for this pathogen siRNA-induced host gene silencing (WEIBERG *et al.* 2013). On the other hand, the important virulence effector Avra10 in the fungal plant pathogen *Blumeria graminis* has been shown to be silenced in barley (*Hordeum vulgare*) and wheat (*Triticum aestivum*) by RNA delivered from the hosts. In this case, an artificial system was constructed in the host to accumulate a reasonable amount of dsRNA or antisense RNA that knocks down this fungal virulence gene during infection

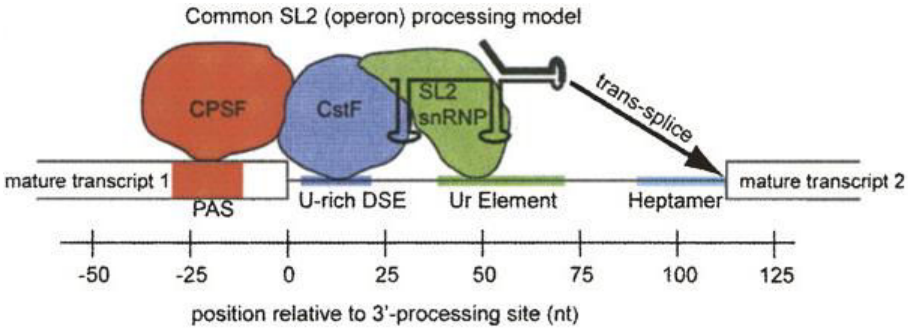


Figure 1.5 Graphical representation of standard internal operon processing, including 3′-processing of the upstream gene, and *trans*-splicing of an SL2 leader to the downstream gene. Copied from (GRABER *et al.* 2007)

(NOWARA *et al.* 2010). As shown by the plant-fungal pathogen interaction, mobile RNA from the host or pathogen can be transferred to its counterpart and silence the targeted gene *in trans*.

We speculated that in our infection model, *C. elegans* and *D. coniospora*, we might be able to feed *D. coniospora* infected worms with RNAi bacteria expressing dsRNA targeting a fungal gene to silence fungal genes potentially involved in virulence.

1.2.1.4 *C. elegans* mRNA trans-splicing and spliced leader 2

As there is no evidence for fungal mRNA inside worm cells, through feeding RNAi bacteria, we can only generate dsRNA or primary siRNA. Since secondary siRNA can make an important contribution to gene silencing (NOWARA *et al.* 2010), we hypothesised that making the 2^{ry} siRNA in worms could increase the possibility of silencing a fungal gene. For this to be possible, one needs first to express the fungal mRNA inside the worm, to allow the generation of the 2^{ry} fungal siRNA. To express the fungal mRNA and visually confirm its expression with a fluorescent tag inside worms, we sought a molecular element that allows the co-expression of a fungal gene and a fluorescent tag in a bi-cistronic manner. As a well-established model organism, several co-expression strategies exist in *C. elegans*, such as the *C. elegans* endogenous internal ribosomal entry site (IRES)-like sequence (LI AND WANG 2012), a recently developed 2A viral peptide technology (AHIER AND JARRIAULT 2014) and the spliced leader 2 (SL2) an endogenous trans-splicing element. As SL2 allows two genes to be expressed separately on the translational level, we choose it as the basis of our fungal gene expression system.

SL2 is a short trans-splice element and in *C. elegans* it is strongly associated with operons. The SL2 associated small nuclear ribonucleoprotein particle (snRNP) binds to the upstream gene as donor of the 5' splicing site and this SL2 will be trans-spliced to the downstream gene behind

the 3' splicing site (Figure 1.5). The SL2 trans-splicing separates the upstream and downstream genes produced under the same promoter and leaves both genes to be translated independently (KUERSTEN *et al.* 1997; LIU *et al.* 2003).

1.2.2 Transformation

1.2.2.1 Polyethylene glycol (PEG) - mediated transformation

At alternative, but more laborious is genetic transformation, basic approach of biological functional study. Delivering and integrating exogenous DNA into the genome of an organism is a way to address biological questions at the molecular genetic level. For eukaryotes, transformation was first performed on several model species such as *Saccharomyces cerevisiae* (BEGGS 1978; HINNEN *et al.* 1978), *Neurospora crassa* (CASE *et al.* 1979) and *Aspergillus nidulans* (BALLANCE *et al.* 1983; YELTON *et al.* 1984). These initial studies set up a polyethylene glycol (PEG) based protoplast transformation protocol, which has been widely adapted to different fungus and plants.

There are three major steps for PEG-mediated transformation: (1) protoplast preparation, (2) transformation of plasmid DNA into protoplasts and (3) selection of transformants on selective media. In general, the protoplasts are produced by removing the cell wall from mycelia or germinating spores by enzymatic digestion. These protoplasts will be stabilized in high salt and sorbitol solution before and after transformation. The PEG will be added to facilitate the DNA uptaking after incubation of high concentrations of DNA and protoplasts. Notably, it is still unclear how PEG triggers DNA uptake into the protoplast (KUWANO *et al.* 2008). The stabilized transformants are selected based on the antibiotic or auxotrophic marker included on the transforming DNA.

Although PEG-mediated transformation is a simple and commonly used method (RUIZ-DIEZ 2002; LIU AND FRIESEN 2012), there are certain drawbacks for this protocol. For instance,

because of generally low transformation efficiency, high concentrations of protoplasts are needed. This can be a problem for slow-growing fungi. Otherwise, in part because of the high DNA concentrations used, integration can occur at multiple loci, making it more difficult to obtain the desired strain. To overcome these drawbacks, several methods have been developed, such as electroporation (CHAKRABORTY AND KAPOOR 1990), *Agrobacterium* mediated transformation (BUNDOCK *et al.* 1995; DE GROOT *et al.* 1998) and biolistics (ARMALEO *et al.* 1990). During my work, I experimented first with PEG-mediated transformation. As described below (0), fortunately, it proved a suitable method for transforming *D. coniospora*.

Turning to methods to construct plasmids for transformation, there is a variety of high-throughput techniques available, such as Gateway™ (HARTLEY *et al.* 2000). This method relies on a site-specific recombination reaction of bacteriophage λ enzymes that enables an insert to be joined with the sequence-specific sites of a donor vector by *in vitro* recombination. However, this method is only efficient with relatively short or simple constructs. Making combinations between different fluorescent tags and genes of interest, for example, can be laborious. In addition, the method would also introduce irrelevant DNA sequences, the end-joining sequence between the insert and vector, into the fungal genome after transformation. In comparison with Gateway™ technology, Gibson Assembly™ can assemble any plasmid components with a minimum requirement of 15 bp of homologous sequence.(GIBSON *et al.* 2009a). It is also commonly used for *in vitro* DNA assembly nowadays.

Gibson and his colleagues developed the technique in 2004 [https://www.neb.com/~/media/NebUs/Files/Feature%20Articles/Gibson_Assembly.pdf]. They discovered a combination of 3 enzymatic activities were able to assemble overlapping DNA fragments efficiently during a project to synthesize the *Mycoplasma genitalium* genome: (a) an

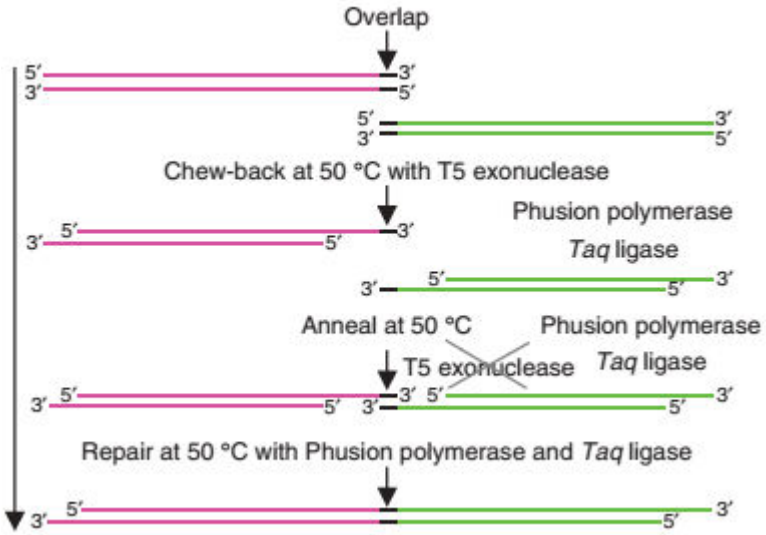


Figure 1.6 One-step isothermal *in vitro* recombination with Gibson Assembly (copied from GIBSON *et al.* 2009b).

exonuclease activity which chews back the ends of DNA fragments and leaves a single strand DNA overhang that can anneal to its complement; (b) DNA polymerase activity is needed to fill the gap on the annealed fragments; (c) DNA ligase activity will be used to seal the resulting nicks on the almost joined fragments after gap filling (GIBSON AND RUSSELLO 2010). Since then, an *in vitro* assembly method was developed based on these 3 enzymatic activities in a two-step thermocycle. This method has been successfully used to assemble the *M. genitalium* genome into 4 large parts (136 - 166 kb) from 101 overlapping DNA cassettes (GIBSON *et al.* 2008). With further improvement, the two-step thermocycle approach was developed into Gibson Assembly, a one-step isothermal assembly method based on 3 enzymes: T5 exonuclease chews 5' ends of DNA fragments, Phusion DNA polymerase fills the gaps and *Taq* DNA ligase seals the nicks (Figure 1.6 GIBSON *et al.* 2009b).

To construct a desired plasmid, a fragment of DNA will be inserted into the linearized vector, but often it is possible that the recirculation could happen with or without DNA insertion. There are several techniques used to increase the selection of positive colonies. For instance, blue/white screening which distinguishes the recombinants from the non-recombinants based on the colour of the colony (MESSING *et al.* 1977); alkaline phosphatase which removes the 5' terminal phosphate on the vector and inhibits self-ligation; the *ccdB* selection system, the toxin encoding gene *ccdB* in the vector will be replaced by the insert which allows the selection of the right recombinants (BERNARD *et al.* 1994).

The *ccdB* gene encodes a topoisomerase toxin which targets the GyrA subunit of DNA gyrase, an essential type II topoisomerase in *Escherichia coli*. The CcdB induces the formation of un-resolvable GyrA-DNA complex that increases plasmid or chromosomal DNA breakage and results in cell death (BERNARD *et al.* 1993). The donor vector containing *ccdB* can only survive in

a *ccdB* resistant *E. coli* strain. After the recombination and transformation only the plasmid with its ccdB replaced by the insert will survive in non-resistant *E. coli* strain (BERNARD *et al.* 1994). The general utility of this method motivated me to produce *ccdB*-containing vectors suitable for use with *D. coniospora*.

1.2.2.2 Vegetative hyphal fusion and *so* gene

Vegetative hyphal fusion (VHF) is a defining feature of filamentous fungi. This has been studied intensively in *N. crassa* (ROCA *et al.* 2005). During the vegetative growth stage, VHF allows fungi to build up multicellular colonies through two processes, conidial anastomosis tube (CAT) fusion and fusion of specialized hyphae. The fungus may benefit from VHF in several ways. Generally, both CAT and hyphal fusion are important for the distribution of nutrients and water, and intrahyphal communication within the fungus (GREGORY 1984; DAVIDSON *et al.* 1996). Particularly, CAT fusion or VHF in general provides a way for horizontal gene transfer between different conidia or strains (DEBETS *et al.* 1994; ROCA *et al.* 2004; ROCA *et al.* 2005). This could be a possible explanation for the high genetic variation among asexual fungi.

Previous studies showed that the *so* gene is essential for VHF among the filamentous fungi (FLEISSNER *et al.* 2005; PRADOS ROSALES AND DI PIETRO 2008). In *N. crassa*, SO is recruited together with MAK-2 to the plasma membranes of two interacting germilings in an oscillatory manner. During CAT fusion, inhibiting MAK-2 function in one CAT causes the stabilization of SO and abolishment of MAK-2 in the partner CAT tip, leading to the failure of CAT fusion (FLEISSNER *et al.* 2009). These data suggest that SO either prevents the localization of MAK-2 or serves as an essential component for VHF signal transduction (LEEDER *et al.* 2011). Interestingly, a study in *Fusarium oxysporum* demonstrated that a *so* mutant (Δ*fso*) is only

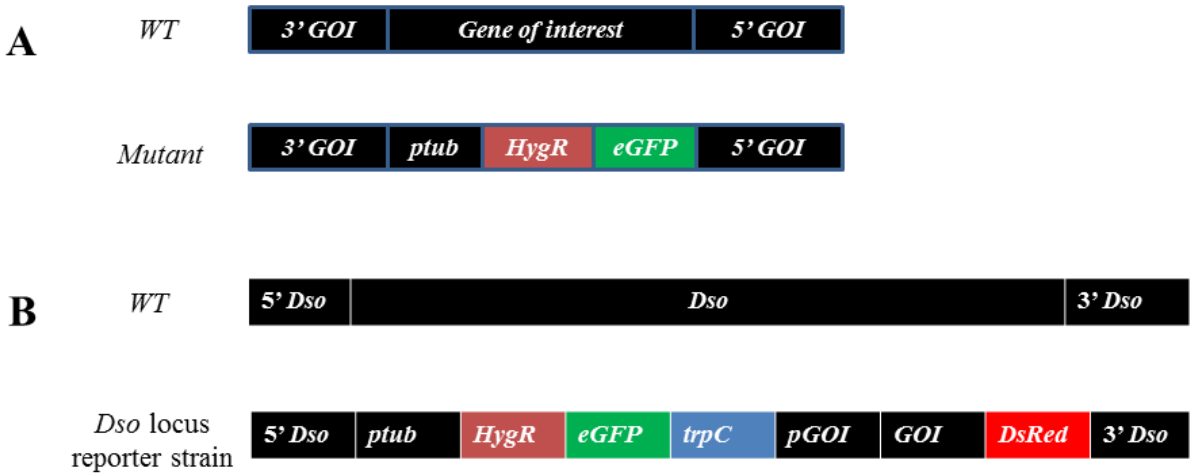


Figure 1.7 Schematic representation of the gene deletion (A) and *Dso* locus c-terminus protein tagging (B) strategy. (A). Hygromycin resistance gene and GFP chimeric cassette driven by *β*-tubulin promoter was used to replace gene of interest. (B). *Dso* was replaced by the gene of interest with its DsRed tag together with Hyg^R and GFP chimeric cassette.

deficient for VHF but not for its virulence (PRADOS ROSALES AND DI PIETRO 2008). Targeting *so* therefore provides a mean of generating mutants with an easily observable phenotype, and prevents fusion of filaments from different colonies.

1.2.2.3 The scheme for reporter or knock-out plasmid

There are two ways to help understanding the function of a specific gene, either by knocking it out or generating a reporter strain to determine its spatiotemporal expression. During this project, I employed a homologous recombination method, replacing the gene of interest with a Hygromycin resistance (Hyg^R) and GFP fluorescent chimeric cassette driven by the β-tubulin promoter (Figure 1.7 A). For reporter strains, I choose to replace the *Dso* gene with the gene of interest and its fluorescent tag together with the GFP- Hyg^R chimeric gene (Figure 1.7 B), as the *Dso* gene is not essential but its mutation is expected to provoke a clear phenotype.

1.3 A *D. coniospora* effector and its possible host target

1.3.1 Fungal effectors and saposins

Host-pathogen interactions can exist in a gene for gene manner. This theory was proposed through the study of plant-pathogen interactions, with the definition of host resistance (R) genes and their counterpart pathogen avirulence (Avr) genes (FLOR 1971). The plant R gene product interacts with the pathogenic Avr virulence effector, so that Avr effector will be disarmed, leading to plant resistance (Figure 1.8). To be able to mount an efficient infection, the pathogen has to secrete proteins into the host cell membrane or cytoplasm and manipulate host immunity by targeting its immune components. These secreted proteins in fungal pathogen are termed effectors. They have been intensively studied in plant fungal pathogens. As detailed above (section 1.1.3), the effectors of plant fungal pathogen can be divided into 2 groups, the apoplastic effectors that are localized in apoplast and host cytoplasmic effectors. Apoplastic effectors are

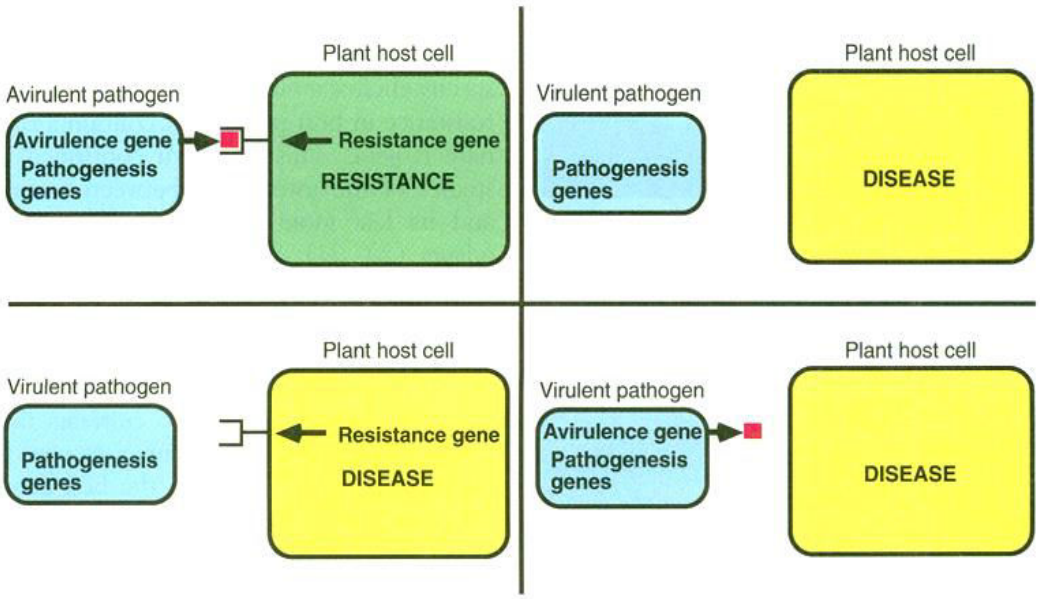


Figure 1.8 Gene-for-gene interactions specify plant disease resistance. (adapted from STASKAWICZ *et al.* 1995)

mostly cell-wall-degrading enzymes (CWDEs) and necrosis inducing proteins. The necrosis-inducing protein triggers cell death by permeabilizing the plant cell (OTTMANN *et al.* 2009). Several cytoplasmic fungal effectors have been reported to be functionally similar to bacterial type III effectors which can suppress host immunity (HOUTERMAN *et al.* 2008; DOEHLEMANN *et al.* 2009).

Whole genome sequencing provides a way of identifying potential fungal effectors. On the basis of predicted genes, one can narrow down the list of candidates on the basis of functional domains and secretion peptides. Since the first fungal pathogen genome published in 2005 (DEAN *et al.* 2005), the number of available genomes has expanded. There are now more than 120 genomes of *Ascomycetes* pathogens in NCBI. The comparison between these genomes can give more insight into the fungal effectors from an evolutionary viewpoint and help one to identify species-specific or conserved effectors from the species of interest (LO PRESTI *et al.* 2015).

To counterattack a pathogen, host organisms can also produce different effectors. As introduced in 1.1.1, *C. elegans* uses several different families of effectors to combat infection, one of which is the saposin-like proteins. Saposins are named from sphingolipid activator proteins due to the fact that they are required for the lysosomal hydrolysis of a variety of sphingolipids (O'BRIEN AND KISHIMOTO 1991). The saposins in mammals are synthesized as a single precursor protein, prosaposin. It contains 2 SapA domain that will be cleaved to release the 4 SapB domain containing saposins, saposin A, B, C and D (PONTING 1994). Among the 4 saposins, saposin B was the first discovered for its role in facilitating the hydrolysis of cerebroside sulfate by activating arylsufatase A (reviewed in LI *et al.* 1985). Similarly, saposin A and C can activate β-glucosylceramidase and β-galactosylceramidase respectively. The function of saposin D, on the other hand, has not yet been determined. Their activator role is probably due

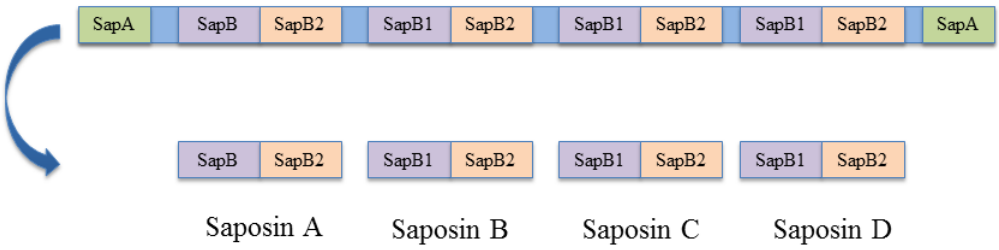


Figure 1.9 Schematic diagram of prosaposin and SapA cleavage.

to their affinity with specific lipid substrates; they can gather these lipids on the cell membrane, so that the degradative enzyme works more efficiently on the concentrated substrates. Surfactant-associated protein B (SP-B) is another protein that has an enzymatic SapB domain activated by the cleavage of a SapA domain. SP-B is a crucial lipid-associated protein in lung surfactant.

SAPLIP stands for saposin-like protein and is a family of SapB domain containing proteins that interact with lipid but have various functions. They are well conserved and can be found from primitive protozoa to mammals, and include protozoan amoebapores and mammalian prosaposins. In *C. elegans*, there are 33 potential proteins with a SapB domain encoded by 28 genes (ROEDER *et al.* 2010). These proteins are specifically classified in the SPP family and named as caenopores in *C. elegans*, as they are functionally and structurally similar to amoebapores, known antimicrobial and pore-forming proteins. Several of these proteins, such as SPP-5, SPP-1 and SPP-12 have been shown to be important for intestinal immunity and have a strong ability to permeabilize microbial cell membranes (ROEDER *et al.* 2010; HOECKENDORF *et al.* 2012). Importantly, several SPP proteins are up-regulated upon *D. coniospora* infection (ENGELMANN *et al.* 2011).

A part of the proprotein, the SapA domain peptide in the C terminus of pro-saposin and the N terminus of pro-SP-B, is important for protein targeting and transport to the lysosome and to secretory granules, respectively (LIN *et al.* 1996; LEFRANCOIS *et al.* 2002). The SapA domain is considered to shield the SapB domain, keeping saposins dormant before the cleavage. Analysis of the *D. coniospora* genome revealed the presence of 2 genes encoding SapA domain-containing proteins. This led to the hypothesis that fungal SapA domains could interact with host saposins

and impair their function. We therefore specifically addressed the interaction between a fungal SapA domain containing protein (G3895*) and the *C. elegans* caenopores (see 0).

1.3.2 Ascorbate peroxidase fingerprinting-based effector discovery

Apart from the bioinformatics analysis of the genome sequence, there are also other methods to fish out pathogen effectors, such as biochemical approaches. Compared to the bioinformatics analysis, biochemical identification is in some ways a more straightforward method, as only functional effectors will be detected. Furthermore, considering that not all effectors possess a clear secretion signal, the biochemical approach will not only allow validation of predicted effectors but also provides an unbiased approach to uncover novel virulence proteins.

In order to identify pathogen effectors in a comprehensive yet specific manner, it is necessary to target and separate pathogen effectors within different compartments of the host cell. Recently, a streptavidin-biotin based spatially restricted enzymatic tagging method has been developed for labelling proteins in specific subcellular locations (RHEE *et al.* 2013). An ascorbate peroxidase (APEX) is used in this approach. APEX can catalyse the transformation of biotin-phenol substrate to biotin-phenoxyl radicals in the presence of H_2O_2, thereby promoting the biotinylation of near-by proteins. These biotinylated proteins can be purified with streptavidin-beads and identified with mass spectrometry (RHEE *et al.* 2013). This approach has been used in the proteomic mapping of human mitochondrial sub-compartments (HUNG *et al.* 2014), the *Drosophila* mitochondria matrix (CHEN *et al.* 2015) and the cilium (MICK *et al.* 2015). In *C. elegans*, the Troemel group adapted this protocol to study *Microsporidia* host-exposed proteins by expressing APEX in intestinal cell nuclei or in the cytoplasm (REINKE *et al.* 2016).

*The gene of *D. coniospora* named by our lab is prefixed with "g", its protein is prefixed with "G"

This method has been further developed for the purpose of studying *D. coniospora* effectors (see 4.4).

In the following section, I present my work that I conducted for my PhD thesis. Given the fact that different subjects are not intuitively related, for each part I give the results together with the discussion and perspectives for future work. When a paper involved several other people, I explain my contribution.

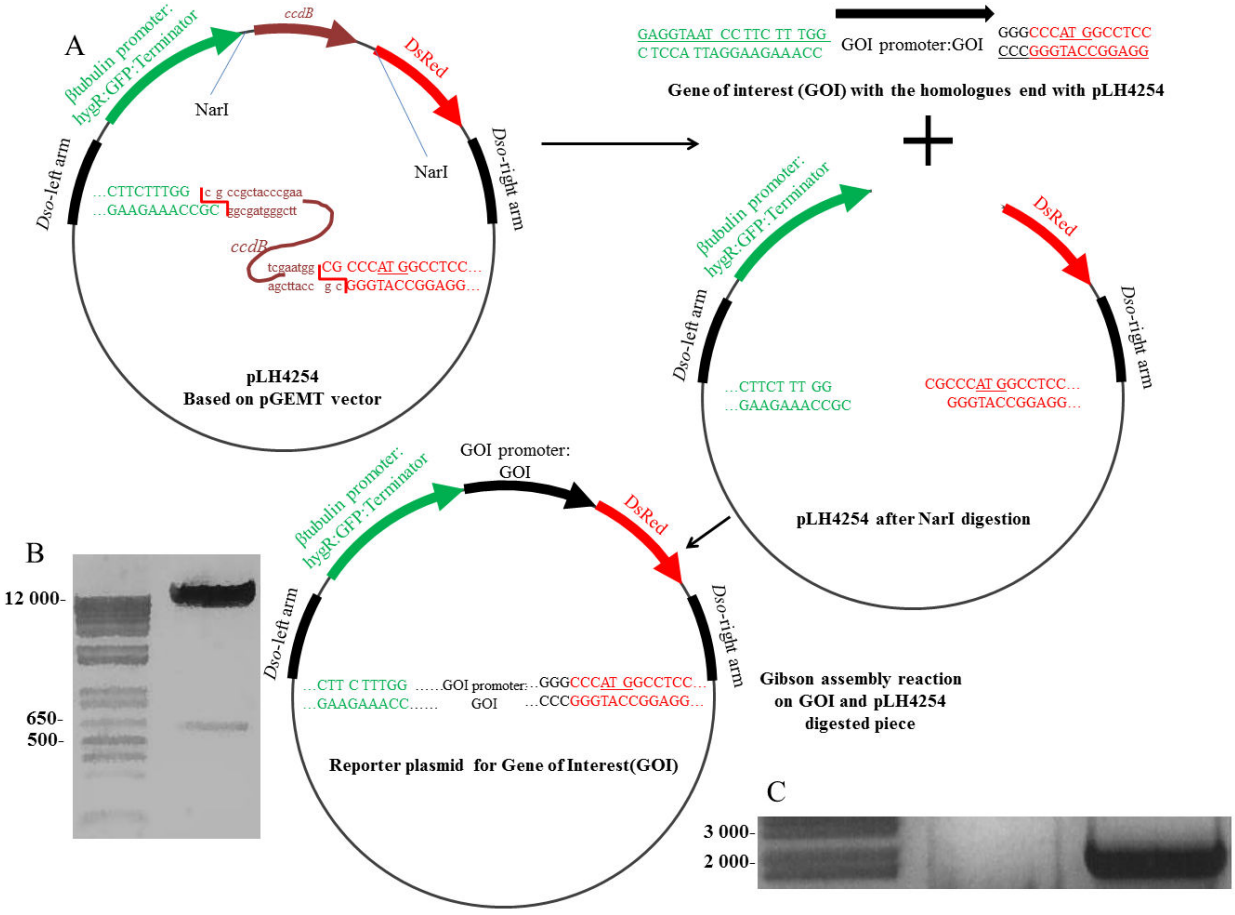


Figure 2.1 Reporter plasmid construction based on pLH4254. (A). Schematic representation of the reporter plasmid cloning strategy. (B). Original plasmid digested with *NarI* restriction enzyme. (C). g4535 and its 5' UTR was amplified with jep2849/jep2850.

Table 2.1 Plasmids constructed based on Gibson Assembly and *ccdB* selection

Plasmid	Gene of interest	Function domain prediction	Reporter of Knock out
pLH4255	g4535	ZnMc MMP like 1, CD04279	Reporter
pFB4256	g1409	Secreted, Pertussis S1, PF02917	Reporter
pFB4257	g2698	Ribosome inactivation protein, PF00161	Reporter
pFB4258	g2819	Enterotoxin A, PF01375	Reporter
pLH4253	g3895/sapA	SapA, PF02199	Knock-out
pFB4259	g4535	ZnMc MMP like 1, CD04279	Knock-out

2.1 Plasmid construction based on Gibson assembly and *ccdB* selection

Two plasmids were constructed with the purpose of making reporter plasmids (pLH4254, Figure 2.1 A), or knock-out plasmids (pLH4252, Figure 2.2 A). For the reporter vector plasmid, two *NarI* restriction enzyme sites harbouring a *ccdB* cassette were inserted. After *NarI* digestion, the plasmid will be separated into two fragments, 0.6 kb and 9 kb respectively, providing an insertion site for the gene of interest and its promoter (Figure 2.1 B). For constructing a reporter plasmid, we assembled the purified DNA of the gene of interest (GOI) together with 1 kb 5' flanking sequence and digested pLH4254 in the Gibson Assembly solution. The resulting plasmid (Figure 2.2 C) was verified by sequencing.

For the knock-out vector, two *ccdB* cassettes are linked by two sets of *NotI* and *ClaI* sites to eGFP and Hyg^R cassettes driven by the beta tubulin promoter. Four fragments will be generated after digestion, two *ccdB* fragments (0.6 kb), the beta tubulin promoter with the eGFP-HygR chimeric gene (3.5 kb) and the vector backbone (3 kb) that will be ligated with the left and right arms of the gene of interest (Figure 2.2 B). Taking the predicted gene g4535 as an example, its left and right arms were amplified by PCR and purified (Figure 2.2 C), and further assembled with the digested pLH4252. The resulting plasmid was verified by sequencing.

Based on this protocol, 6 plasmids were made for 5 genes (Table 2.2) including g3895, predicted to encode a SapA domain protein.

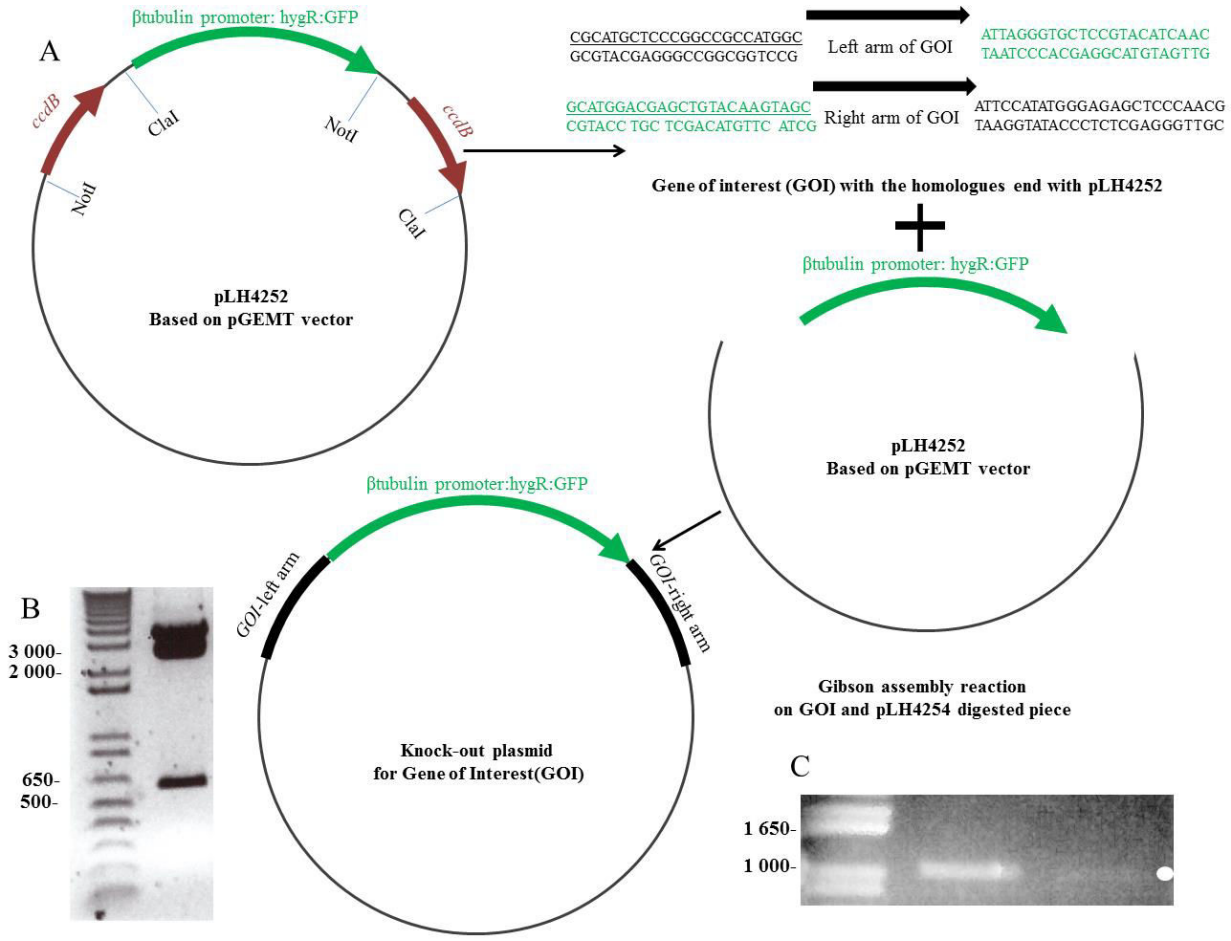


Figure 2.2 Knock-out plasmid construction pLH4252. (A). Schematic representation of the reporter plasmid cloning strategy. (B). Original plasmid digested with *NarI* restriction enzyme. (C). Left arm and right arm of g4535 were amplified with jep2878/jpe2879 and jep2880/jep2881. White point was marking the right arm band.

2.2 Article 1

Polyethylene glycol-mediated transformation of *Drechmeria coniospora*

Le D. He and Jonathan J. Ewbank

Bio-protocol. Under review

Polyethylene glycol-mediated transformation of *Drechmeria coniospora*

Le D. He and Jonathan J. Ewbank

Centre d'Immunologie de Marseille-Luminy, Aix Marseille Université, Inserm, CNRS, Marseille, France

Abstract

Drechmeria coniospora is a nematophagous fungus and potential biocontrol agent. Together with its natural host, *Caenorhabditis elegans*, it is used to study host-pathogen interactions. Here, we report a polyethylene glycol-mediated transformation method [1,2] for this fungus. The protocol can be used to generate both knock-in or knock-out strains [3].

Materials and reagents

Escherichia coli strain OP50
Caenorhabditis elegans strain N2
Drechmeria coniospora recipient strain ATCC 96282
Ampicillin sodium salt (Sigma-Aldrich, catalog number: A9518)
Gentamicin solution (Sigma-Aldrich, catalog number: G1272)
Hygromycin B (Thermo Fisher scientific, catalog number: 10687010)
Bacto™ agar (BD Biosciences, catalog number: 214010)
Bacto™ peptone (BD Biosciences, catalog number: 211677)
Yeast extract (Douchefa Biochemie, catalog number: Y1333.0500)
D-sorbitol (Sigma-Aldrich, catalog number: S3889-1KG)
Calcium chloride (Sigma-Aldrich, catalog number: C1016)
Sodium Chloride (VWR, catalog number: 27810.295)
Magnesium sulfate (Sigma-Aldrich, catalog number: M7506)
di-Potassium hydrogen orthophosphate (VWR, catalog number: C26931.263)
Potassium hydrogen sulphate (VWR, catalog number: C27011.294)
Tris/HCl (Sigma-Aldrich, catalog number: T6666)
Polyethylene glycol (PEG 3350) (Sigma-Aldrich, catalog number: P4338)
Caylase C4 (CAYLA, catalog number: caseC4-5)
Aurintricarboxylic acid (Merck Millipore, catalog number: 189400)
NGM (see Recipes)
NGMY (see Recipes)
NGMS (see Recipes)
CaCl 1700 (see Recipes)
STC (see Recipes)

Equipment

Falcon™ 50 ml conical tube (Corning Inc., catalog number: 1443222)
Falcon™ 15 ml conical tube (Corning Inc., catalog number: 1495953)
Falcon™ 14 ml Polystyrene round bottom tube (Corning Inc., catalog number: 352051)
Pasteur pipettes (VWR, catalog number: 6121701)
Syringe (Terumo, catalog number: SS+20ES1)
Aluminum Foil (Fisher Scientific, catalog number: 01213102)
Acrodisc® syringe Filters (Pall Life Sciences, catalog number: 4187)
Pertri dish (Greiner Bio-one, catalog number: 633185)
Miracloth (Millipore, catalog number: 475855-1R)
Counting chambers (Burker, Tiefe 0.1000m, 0.0025mm^2)
Spatula

Funnel
Centrifuge (Eppendorf, catalog number: 5810R)
HERAsafe KS safety cabinet (Thermo scientific, catalog number: 51022751)
Balancer
Stereomicroscope (Leica MZ 16 F)
Rotary shaker for culture flasks (New Brunswick Scientific, catalog number: innova4080)

Procedure

Mycelium preparation

1. Thaw a frozen stock of fungal mycelia and spread on NGM plates. Culture at 25 °C for 1 week. Using a stereomicroscope, confirm appearance of spores. If no spores are visible, continue culture and check daily.
2. Using a bent sterile glass Pasteur pipette, gently rub the fungal colonies to dislodge the spores. Wash the spores from the plates, using 1 ml of 50 mM NaCl (Figure 1A & B) and collect in an Eppendorf tube.
3. Spread the spores (20-100 µl) on a 9 cm NGM-AG plate (Figure 1C) and allow to dry in a safety cabinet. Add 2000-3000 N2 worms that are synchronized at the L4 stage [4]. Incubate at 25 °C for 10 days.
4. Wash off spores as in step 2. The spores collected from one NGM-AG plate can be spread on 4 fresh NGM plates seeded with OP50 [4]. If necessary, allow the solution to dry.

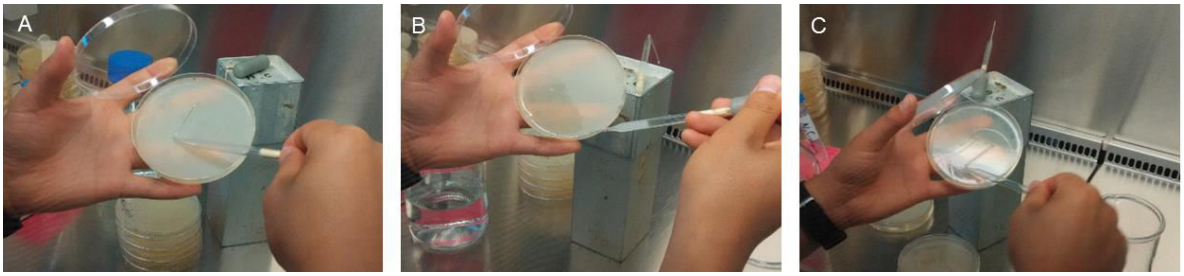


Figure 1. Preparation of infection plates. **A, B**. Washing spores from NGM-AG plate. **C**. Spreading spore solution on an NGM OP50 plate.

5. Transfer 4,000-5,000 L4 or young adult N2 worm onto each spore-OP50 plate, and incubate for 24h.
6. Wash the infected worms from each plate with 1.5 ml 50mM NaCl into 15 ml Falcon tube and centrifuge at 2000 rpm for 1 min; carefully remove the supernatant and divide the worm pellet between three 9 cm NGMY plates. Gently spread the pellet across the surface of the plate. If necessary, allow to dry in a safety cabinet before incubating at 25 °C for 1 day. Most adult worms should be dead by this time.
7. Add 1-2 ml of 50 mM NaCl to each plate, and using a bent sterile glass Pasteur pipette, rub to resuspend the dead worms. Collect the liquid from several plates in a 15 ml Falcon tube; transfer the solution into 250 ml NGMY liquid, culture on a shaking incubator at 170 rpm for 36-40 h.

Protoplasting

8. Pour the liquid culture into a funnel lined with Miracloth to filter the mycelia; wash with 100 ml CaCl 1700 buffer, use a spatula to transfer the mycelia into a 50 ml Falcon tube (Figure 2). Keep the used funnel and Miracloth in a sterile environment for later use.

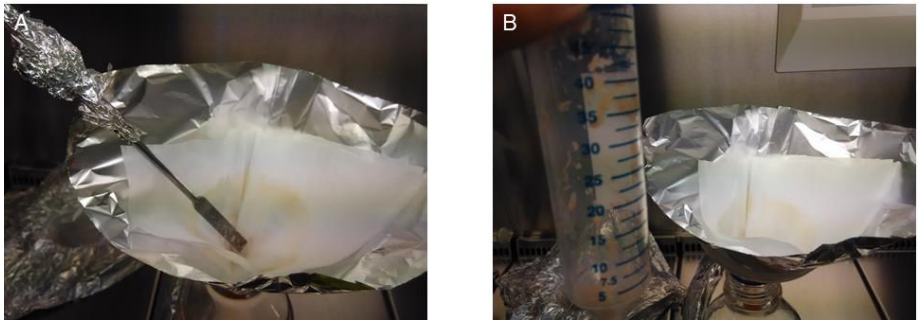


Figure 2. Collecting mycelium with Miracloth

9. Weigh the mycelia and resuspend it in newly prepared lysis buffer; for every 1 g mycelia add 10 ml lysis buffer. If the volume exceeds 30 ml, use a sterile, cloth-plugged Erlenmeyer flask, otherwise a 50 ml Falcon tube.
10. Incubate the mycelia solution at 30 °C with 60 rpm for 1-1.5 h. After 1 h, at regular intervals, remove a small aliquot and check the solution under the microscope. If short mycelial fragments appear, stop the culture (Fig. 3).

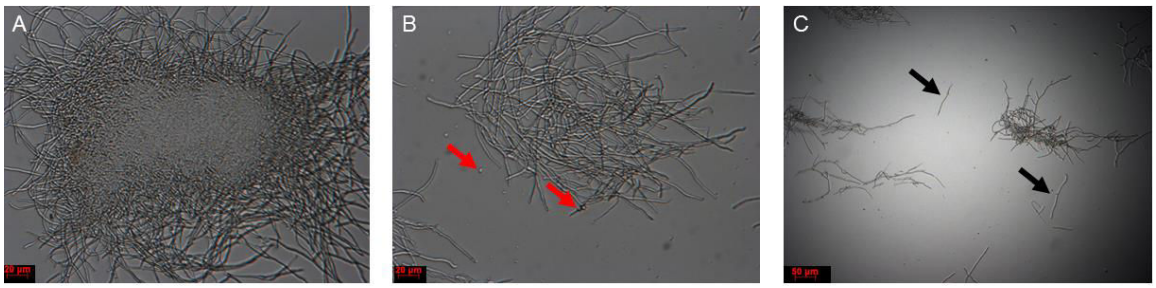


Figure 3. Mycelium solution before and after lysis buffer digestion. **A.** Colony before treatment; **B**, colony after 1 h of treatment, with protoplasts (red arrows). **C**. Mycelial fragments after treatment (black arrows).

11. Place the funnel together with the Miracloth used at step 8 onto a 50 ml Falcon tube. Filter the mycelia and protoplasts in lysis buffer, retain the filtrate.
12. Add an equal volume of ice-cold STC buffer in protoplast solution, mix gently and leave on ice for 10 min.
13. Centrifuge the mixture at 4 °C, 2500 rpm for 15 min.
14. Carefully discard the supernatant; gently resuspend the pellet with 25 ml ice-cold STC buffer; centrifuge the mixture at 4 °C, 3000 rpm for 10 min.
15. Repeat step 14 one more time.
16. Discard the supernatant, add 300 µl ice-cold STC buffer to resuspend the protoplasts; remove a small aliquot and count the protoplasts using a haemocytometer. Dilute the protoplasts to a final concentration to 10^8 protoplasts/ml using ice-cold STC buffer. Then, stock the solution on ice or at 4 °C; it can be used for transformation for a maximum of 7 days.

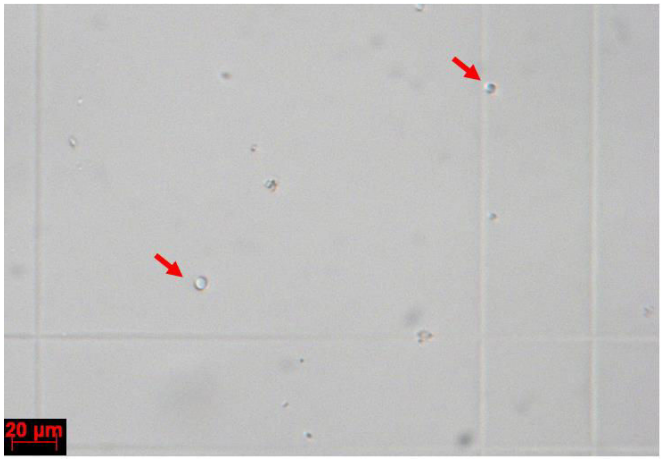


Figure 4. Protoplast on heamocytometer

Transformation

17. Add 1 µl ATA onto the wall of a round-bottom 14 ml Falcon tube, mix with 10 µg or more of plasmid DNA (in a volume of around 10 µl), then mix with 150 µl of protoplast solution by gently pipetting; incubate the mixture for 40 min at room temperature.
18. Using a Pipetman P1000, add 6 drops of PEG1700 buffer, mix by gently shaking; add another 850 µl PEG buffer, gently shake again, incubate for another 20 min.
19. Add 10 ml of room temperature STC buffer, centrifuge 10 min at 3000 rpm at room temperature.
20. Discard the supernatant carefully; resuspend the protoplasts with 300 µl STC at room temperature.
21. Add 150 µl of protoplast solution to a 9 cm plate (see note 2), spread gently with sterile bent glass Pasteur pipette; culture at 25 °C for 3 weeks,
22. Check any colonies for successful transformation, by PCR or under a fluorescent stereomicroscope when using a transformation construct that includes a fluorescent reporter gene.

Notes

1. Unless specified, temperature is always at 25 °C.
2. Except OP50 NGM plates, all the other plates are with ampicillin and gentamicin. When transformation constructs include a hygromycin-resistance cassette, NGMS plates are supplemented with 25 µg/ml hygromycin.
3. For step 14, first use 1 ml of cold STC to resuspend the pellet, and then add the remaining 24 ml, otherwise the pellet will be hard to dissolve.
4. For step 9, mycelia fragments can give rise to colonies that are quite resistant to hygromycin selection; avoid fragmentation as much as possible.
5. Autoclave the Miracloth (folded in half twice), in an aluminum foil bag with a cut corner, so that condensation does not accumulate and the cloth is dry after autoclaving.

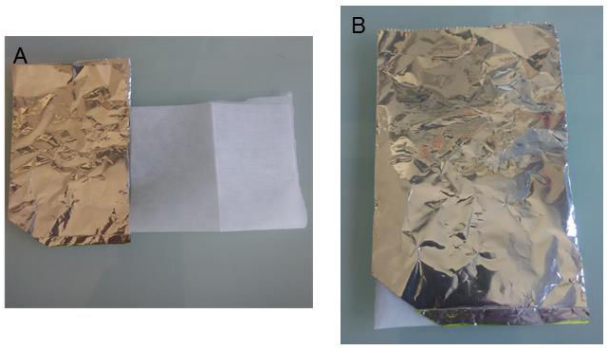


Figure 5 Miracloth preparation

Recipes

Medium

NGM (1L, autoclave)

Bacto™ agar	20 g
Bacto™ peptone	2.5 g
NaCl	3 g
5 mg/ml cholesterol in ethanol	1 ml
After autoclave	
1 M $MgSO_4$ (sterilized)	1 ml
1 M $CaCl_2$ (sterilized)	1 ml
1 M phosphate buffer pH 6.0 (sterilized)	1 ml

Phosphate buffer: 108.3 g KH_2PO_4, 35.6 g K_2HPO_4, H_2O to 1 l

NGM-AG: NGM with 100 µg/ml ampicillin and 50 µg/ml gentamicin

NGMY: NGM plus yeast extract 20g/l

NGMY liquid: NGMY without agar

NGMS: NGM-AG plus d-sorbitol

Solution

ATA 15 mM solution (autoclave)

NaCl 50 mM solution (autoclave)

CaCl 1700 (500ml, autoclave):

$CaCl_2$	15.0 g
NaCl	17.5 g

Lysis buffer (prepare fresh every time, filter with syringe and film): CaCl 1700: C4 Caylase = 10 ml: 100mg

STC 1700 (500ml, autoclave):

Sorbitol	109.3 g
Tris/HCl	5 ml
CaCl2	2.8 g
NaCl	1 g
PEG 1700 (100ml, autoclave):	
PEG 4000	60 g
Tris/HCl	1 ml
$CaCl.2H_2O$	0.74 g

* Tris/HCl solution stored at 1 M pH 7.5	60.6 g/500 ml	M=121.2

Acknowledgments

We thank Eric Record for the help of setting up the protocol. The Jonathan Ewbank lab members' input and suggestions.

References

[1] B.G. Turgeon, B. Condon, J. Liu, N. Zhang, Protoplast Transformation of Filamentous Fungi, in: A. Sharon (Ed.) Molecular and cell biology methods for fungi, Humana, 2010.
[2] H. Ochman, A.S. Gerber, D.L. Hartl, Genetics, 120 (1988) 621-623.
[3] Lebrigand K, He le D, Thakur N, Arguel MJ, Polanowska J, Henrissat B, Record, E, Magdelenat, G, Barbe, V, Raffaele, S, Barbry, P, Ewbank, J. J, . PLoS Genet. 2016;12(5):e1006017.
[4] Stiernagle T. Maintenance of C. elegans. The *C. elegans* Research Community ed.; 2006. Available from: http://www.wormbook.org.

Table 2.2 Transgenic strains obtained with the protocol of Article 1

Lab code	**Plasmid used**	**Genotype**
JEM763	pPK2-hphgfp	P*gpd*::HPH::eGFP::*trpC*
JEM818	pLH4241	5' *Dso*::Ptub:: HPH::eGFP::3' *Dso*
JEM819	pLH4237	P*tub*:: HPH::eGFP::*trpC*
JEM907	pLH4244	5' *Dso*::P*tub*::eGFP::hph::trpC::5'flank *sapA*::*sapA*::DsRed::3' *Dso*
JEM931	pLH4238	P*gpdA*::HPH::eGFP::*trpC*

2.2.1 Additional information: transgenic strains list

Even though this protocol proved to be a success for D. coniospora transformation, as we obtained 5 transgenic strains (Table 2.2), we still need to improve it, as often than not we got no transformants.

2.3 Discussion and perspective

PEG-mediated protoplast transformation has been developed as a method for nearly 40 years and it has been proved to be useful in a wide range of fungi (LIU AND FRIESEN 2012). In our lab, we have adapted this protocol successfully to generate 5 different *D. coniospora* strains, including that contain random insertions of a plasmid, either encoding eGFP or DsRed, and two knock-in strains, namely the *Dso* and *sapA* reporter strains (see section 0). Nevertheless, there are certain drawbacks for using this protocol in *D. coniospora*. Indeed, although we have had 5 successes, numerous other attempts at transformation have failed for unknown reasons. Unfortunately, the procedure is time-consuming and complex. The preparation of mycelia involves fungal growth in 3 different media for 5 days and a minimum of 16 days incubation after transformation. The steps of infection and transfer are laborious and include a substantial risk for contamination. Nevertheless, preliminary experiments suggest that this infection procedure cannot be avoided by directly using infection-free vegetative mycelia cultivated *in vitro*, as these were associated with lower transformation efficiency when compared to infection-derived mycelia. This may be due to the rapid cell division that occurs during infection, leading to the production of massive numbers of spores (DIJKSTERHUIS *et al.* 1991). This could make the protoplast chromosomes more accessible for plasmid DNA. The long waiting time after transformation is due to *D. coniospora*'s obligatory pathogenic nature, as it grows very slowly in artificial media. It has been recently reported that adding animal liver and kidney in the medium can facilitate

D. coniospora vegetative growth and decrease its life cycle to 8 days (ZHANG *et al.* 2015). This growth rate, however, is similar to that seen with NGMY medium in our protocol. With the intention of mimicking the host environment, Lohmann *et al.* showed that addition of phospholipids to standard media increased the mycelial growth of *D. coniospora* and its sporulation (LOHMANN *et al.* 1989). Phospholipids or other compounds that could be provided by the host could be tested in the future to identify factors that enhance fungal growth *in vitro*.

Over the last 4 decades, plenty of tools have been developed for genetic manipulation of fungi. The low frequency of homologous recombination, however, remains a limit for successful transformation. To increase homologous recombination, the recently described CRISPR/Cas9 system was introduced to several filamentous fungi and proved to be successful (LIU *et al.* 2015; NODVIG *et al.* 2015; KATAYAMA *et al.* 2016). CRISPR/Cas9 is a system with two main components, the endonuclease Cas9 protein and guide RNA (gRNA); the gRNA guides the Cas9 protein to bind to the targeted DNA and make double strands break (DSB) which increase the chances of homologous recombination.

Regardless of homologous recombination or random insertion, *D. coniospora* has a relatively low transformation frequency compare to the other fungi, such as *N. crassa* and *A. niger*, when applying the conventional PEG-mediated transformation protocol. CRISPR-Cas9 is an interesting tool for the future testing, as the increasing DSB frequency does not just promote homologous recombination but also the transformation frequency. One possible approach would be to apply the system in a two-step way as reported in *Trichoderma reesei* (LIU *et al.* 2015): firstly, construct a Cas9-expressing strain by integrating the Cas9 gene into the fungal genome under the control of an appropriate promoter; and then, transform protoplasts from the Cas9 strain with gRNA and the donor DNA fragment. Given the relative poor success we have had in

transforming *D. coniospora*, investing in this alternative approach seems warranted if the laboratory plans to continue its investigation of fungal virulence strategies.

Table 3.1 Key RNAi proteins in *D. coniospora*

N. crassa	***D. coniospora***	**Identity%***	**E value***
QDE-1, RNA-dependent RNA polymerase Accession: EAA29811.1	G6947	44	0.0
QDE-2, Argonaute protein Accession: ESA42123.1	G8034	29	4e-59
DCL-1, Dicer-like protein Accession: XP_961898.1	G4625	43	0.0
DCL-2, Dicer-like protein Accession: EAA34302.3	G2283	37	0.0

* Result obtained by using blastp suite-2sequences tool in BLAST.

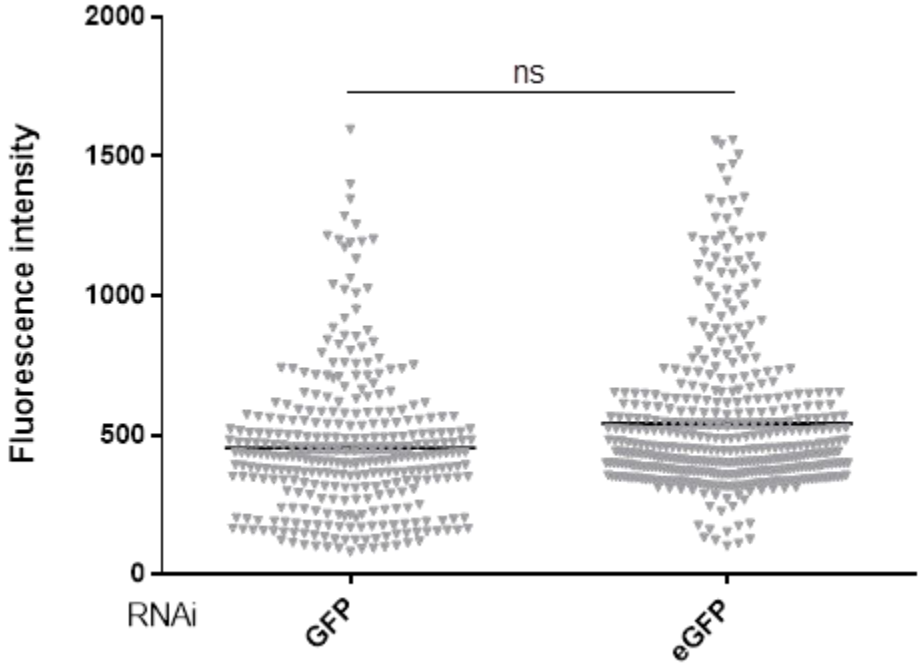


Figure 3.1 RNAi feeding on *rde-1* strain. Worms were fed with GFP or eGFP RNAi bacteria; fluorescence intensity measured after 48h of infection; each dot represents a worm, and its fluorescent intensity is equal to total GFP divided by length; "ns" is not significant, F test was used for the comparison. For each sample, more than 300 worms were used for analysis.

3.1 Primary small RNA for cross-species RNAi

Although the RNAi pathway is generally well conserved among fungi, surprisingly, it is absent in several species (BILLMYRE *et al.* 2013). To verify the presence of RNAi pathway in *D. coniospora*, we examined the existence of the essential RNAi proteins. In *N. crassa*, there are four essential components: the RNA-dependent RNA polymerase QDE-1, the Argonaute protein QDE-2, and the Dicer-like proteins DCL-1 and DCL-2. When we BLASTed these proteins against the set of predicted *D. coniospora* proteins, we found homologues in *D. coniospora* for each of them, with 29-44% protein similarity (Table 3.1).

The existence of these homologues strongly supports the existence of the functional RNAi machinery in *D. coniospora*. As reported for several phytopathogenic fungi, the host derived dsRNA or ssRNA can enter the pathogen and go through its RNAi machinery to silence fungal genes during infection (NOWARA *et al.* 2010; WANG *et al.* 2016). As a proof of principle, we attempted to silence eGFP expression in *D. coniospora* (JEM818) by using the dsRNA or ssRNA targeting eGFP from worm generated by RNAi feeding. An RNAi feeding bacteria strain [eGFP(RNAi)] was generated on this purpose by inserting part of the eGFP gene into the feeding plasmid in HT115 bacteria. It should be noted that the gene of GFP and eGFP are very different (no contiguous stretches of > 5 identical bases), so they can be targeted by RNAi independently.

As introduced before, dsRNA cleavage and production of primary, secondary siRNA producing is dependent on RDE-1 during RNAi feeding in *C. elegans*. Based on this, we used *rde-1* mutant to determine the capability of worm dsRNA for silencing fungal gene. Synchronized adult *rde-1* worms were infected with JEM818 for 48 h and the fluorescent intensity of the infected worm was analysed with Biosorter. The GFP RNAi feeding bacteria [GFP(RNAi)] which we used as a control for RNAi feeding does not targeting eGFP gene. We

found that there was no significant difference between GFP(RNAi) and eGFP(RNAi) fed *rde-1* worms. Which means the eGFP targeting dsRNA from the host did not silence the fungal eGFP expression in our model (Figure 3.1).

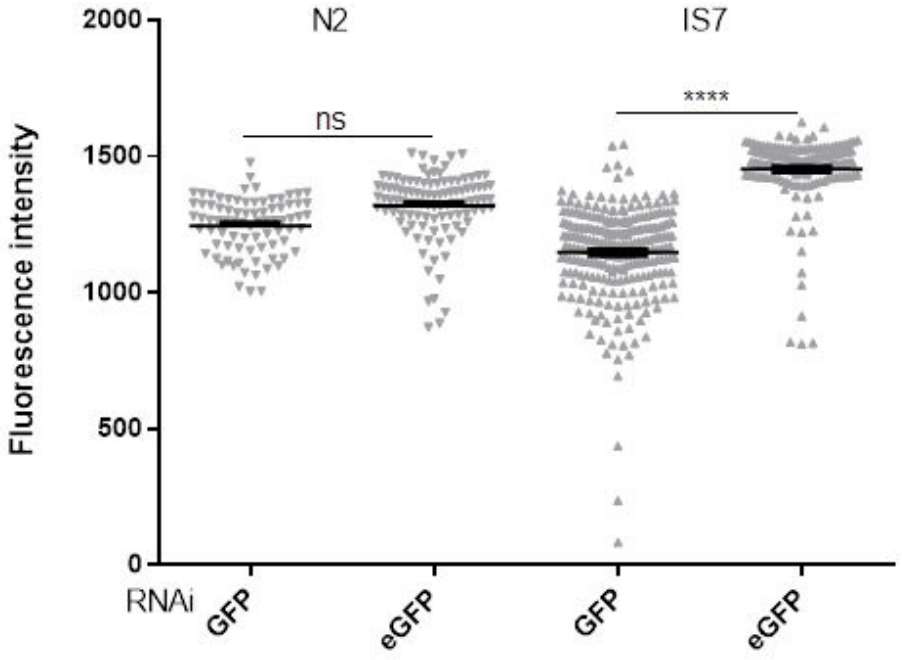


Figure 3.2 RNAi feeding on N2 and IG274 strain. Worms were fed with GFP or eGFP RNAi bacteria; N2 was infected by JEM818 for 48 h; each dot represents a worm, and its fluorescent intensity was total GFP divided by length; the difference between GFP(RNAi) and eGFP(RNAi) on N2 GFP intensity was not significant (ns =not significant, F test); the IG274 (Is7) GFP intensity was significantly different from GFP(RNAi) to eGFP(RNAi) (***, p value <0.0001, F test). Each sample more than 80 worms were used for analysis.

Next, we used WT worm (N2 strain) to address the ability of dsRNA and primary siRNA in combination on their capacity for trans-species RNAi. In this experiment, we used non-infected IG274 (Is7) which express GFP but not eGFP as a control to ensure the RNAi feeding was working. As expected, the green intensity was significantly lower in GFP(RNAi) sample compared to the eGFP(RNAi) sample (Figure 3.2). After 48 h of infection with JEM818, we

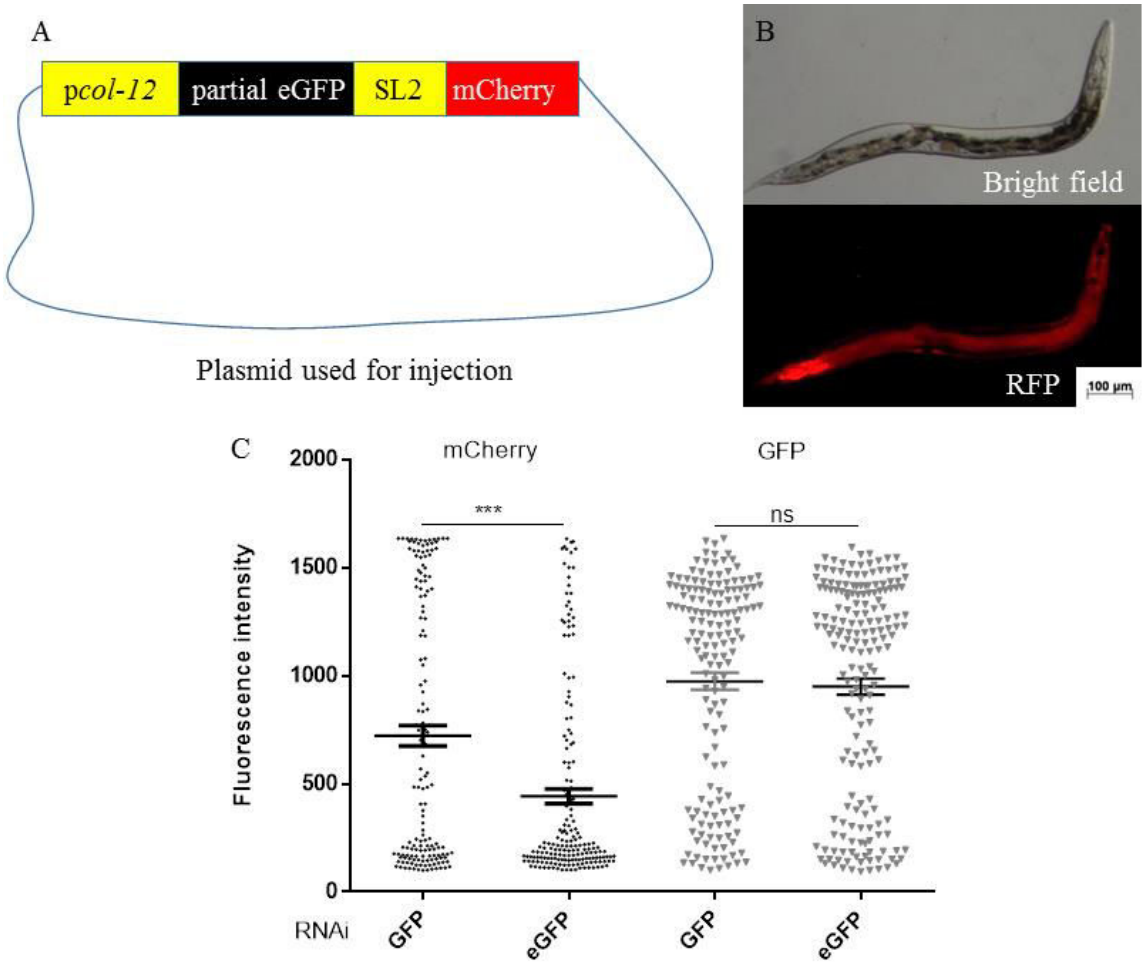


Figure 3.3 Phenotype of strain IG1602 and RNAi feeding on this strain. (A). Plasmid used for generating IG1602. (B). mCherry marker was expressed in the epidermis of IG1602. (C). IG1602 was fed with eGFP or GFP RNAi bacteria and fluorescent intensity was measured after 48 h of JEM818 infection; each dot represents a worm, and its fluorescent intensity was total fluorescent divided with its length; the mCherry intensity was significantly different from GFP(RNAi) to eGFP(RNAi) (***, p value 0.0003, F test); the difference between GFP(RNAi) and eGFP(RNAi) on GFP intensity was not significant (ns =not significant, F test). For each sample, more than 150 worms were used for analysis.

also found there was no difference between GFP(RNAi) and eGFP(RNAi) fed N2 worms with regards green fluorescent intensity. We conclude that neither dsRNA nor primary siRNA in worms was sufficient to silence the fungal eGFP in infected worm.

3.2 Secondary small RNA for cross-species RNAi

To test the capability of secondary (2^{ry}) siRNA for cross-species silencing, we also generated a worm strain (IG1602) which should produce secondary (2^{ry}) eGFP siRNA upon RNAi feeding. In this strain, a bicistronic construct with part of the eGFP gene and the complete mCherry gene is expressed under the *col-12* promoter (Figure 3.3 A). We observed the mCherry (red) expression in the worm epidermis (Figure 3.3 B), indicating that the partial eGFP mRNA was also transcribed.

We measured the red and green fluorescent intensity of the strain IG1602 with the Biosorter after 48 h of infection with the eGFP-expressing fungal strain JEM818. As expected, we observed that the red fluorescent intensity decreased following eGFP(RNAi) compared to GFP(RNAi) (Figure 3.3 C). This decrease presumably reflected the production of 2^{ry} siRNA, as mCherry mRNA was translated from an independent mRNA, so its expression could not have been silenced at the level of translation by dsRNA or primary siRNA. However, we did not observe any difference of GFP intensity in GFP(RNAi) or eGFP(RNAi) treated-sample (Figure 3.3 C).

3.3 Discussion and perspective

Certain types of sRNAs such as dsRNA or duplex sRNA provoke amplification of RNAi , as demonstrated in plants and *C. elegans* (SARKIES AND MISKA 2014). Additionally, ssRNA can be a mobile agent for trans-species RNAi for phytopathogenic fungus (NOWARA *et al.* 2010; WANG *et al.* 2016). We tested the potential capability of dsRNA, primary siRNA or secondary siRNA

for cross-species silencing from *C. elegans* to *D. coniospora*. Our results show that these 3 types of sRNAs from the worm were not sufficient to trigger RNAi in *D. coniospora*. We provide evidence to show that the abundance of sRNA is probably not responsible for this deficiency, as the amount of sRNA is sufficient to trigger RNAi and down regulate mCherry expression in IG1602. However, it is also possible that more time is needed to propagate the RNAi signal inside the host, rather than just 48 h of infection. As the host usually dies after 48 h of infection, in the future, we can try to prolong the duration of the infection by using a more resistant host or an attenuated pathogen strain.

Apart from the perspective of sRNA, it is also possible that *D. coniospora* has a relatively low capability of uptaking sRNA from the external environment. In *C. elegans*, the sRNA can be transferred through vesicles or extracellular fluids via different proteins, the most well characterized proteins being SID-1 and SID-2 (SARKIES AND MISKA 2014). These two proteins, however, have a preference for delivering dsRNA; homologues are absent from fungi, including *D. coniospora*. We cannot exclude the possibility that another type of sRNA transferring machinery is involved in fungus, as it was recently demonstrated that both dsRNA and ssRNA can be delivered into *B. cinerea, B. graminis* and *Verticillium dahlia* from their plant host and silence the fungal Dicer encoding genes or other virulence genes. Nevertheless the mechanism of how these sRNA are transferred from host to pathogen remains unknown (NOWARA *et al.* 2010; WANG *et al.* 2016). It was previously demonstrated that SID-1 can be expressed in *Drosophila melanogaster* S1 cell and form dsRNA uptaking channels (FEINBERG AND HUNTER 2003). Due to the fact that genetic manipulation of the fungus is feasible, we could conceivably enhance its ability for sRNA uptake by heterologously expressing some known sRNA trafficking channels, such as SID-1 or SID-2.

Establishing a cross-species RNAi system in our infection model could bring many advantages for our study. For instance, such a system would allow us to knock down a gene in a less tedious and time consuming way. Furthermore, this system would be able to help addressing the importance of a gene exclusively during infection. Last but not least, it would also be possible to develop this protocol to screen fungal virulence genes in a high through-put manner.

CHAPTER 4. *D. CONIOSPORA* EFFECTORS AND ITS GENOMIC SEQUENCES

4.1 Article 2

Comparative Genomic Analysis of *Drechmeria coniospora* Reveals Core and Specific Genetic Requirements for Fungal Endoparasitism of Nematodes

Lebrigand K, He Le D, Thakur N, Arguel MJ, Polanowska J, Henrissat B, Record E, Magdelenat G, Barbe V, Raffaele S, Barbry P, Ewbank JJ

PLoS Genetetics. 2016 May 6;12(5):e1006017. doi: 10.1371/journal.pgen.1006017.

I performed the experiment and contributed equally with Lebrigand K to this paper.

PLOS | GENETICS

RESEARCH ARTICLE

Comparative Genomic Analysis of *Drechmeria coniospora* Reveals Core and Specific Genetic Requirements for Fungal Endoparasitism of Nematodes

Kevin Lebrigand[1☯], Le D. He[2☯], Nishant Thakur[2], Marie-Jeanne Arguel[1], Jolanta Polanowska[2], Bernard Henrissat[3,4,5], Eric Record[6,7], Ghislaine Magdelenat[8], Valérie Barbe[8], Sylvain Raffaele[9,10], Pascal Barbry[1]*, Jonathan J. Ewbank[2]*

1 CNRS and University Nice Sophia Antipolis, Institute of Molecular and Cellular Pharmacology, Sophia Antipolis, France, **2** Centre d'Immunologie de Marseille-Luminy, Aix Marseille Université UM2, Inserm, U1104, CNRS UMR7280, Marseille, France, **3** CNRS UMR 7257, Aix-Marseille University, Marseille, France, **4** INRA, USC 1408 AFMB, Marseille, France, **5** Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia, **6** INRA, UMR1163 Biodiversité et Biotechnologie Fongiques, Aix-Marseille Université, Polytech Marseille, CP 925, Marseille, France, **7** Aix-Marseille Université, UMR1163 Biodiversité et Biotechnologie Fongiques, Faculté des Sciences de Luminy-Polytech, CP 925, Marseille, France, **8** Commissariat à l'Energie Atomique, Institut de Génomique, Génoscope, Laboratoire de Biologie Moleculaire pour l'Etude des Génomes (LBioMEG), Evry, France, **9** INRA, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR441, Castanet Tolosan, France, **10** CNRS, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR2594, Castanet Tolosan, France

☯ These authors contributed equally to this work.
* barbry@ipmc.cnrs.fr (PB); ewbank@ciml.univ-mrs.fr (JJE)

Citation: Lebrigand K, He LD, Thakur N, Arguel M-J, Polanowska J, Henrissat B, et al. (2016) Comparative Genomic Analysis of *Drechmeria coniospora* Reveals Core and Specific Genetic Requirements for Fungal Endoparasitism of Nematodes. PLoS Genet 12(5): e1006017. doi:10.1371/journal.pgen.1006017

Editor: Paul M. Richardson, MicroTrek Incorporated, UNITED STATES

Received: June 19, 2015

Accepted: April 8, 2016

Published: May 6, 2016

Data Availability Statement: All relevant data are either within the paper and its Supporting Information files or available from the NCBI; MIseq reads used for de novo genome assembly and Hiseq reads used for RNAseq de novo have been deposited at SRA under accession numbers SRX883538 (http://www.ncbi.nlm.nih.gov/sra/?term=SRX883538) and SRX969055 (http://www.ncbi.nlm.nih.gov/sra/?term=SRX969055), respectively. This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JYHR00000000 (http://www.ncbi.nlm.

Abstract

Drechmeria coniospora is an obligate fungal pathogen that infects nematodes via the adhesion of specialized spores to the host cuticle. *D*. *coniospora* is frequently found associated with *Caenorhabditis elegans* in environmental samples. It is used in the study of the nematode's response to fungal infection. Full understanding of this bi-partite interaction requires knowledge of the pathogen's genome, analysis of its gene expression program and a capacity for genetic engineering. The acquisition of all three is reported here. A phylogenetic analysis placed *D*. *coniospora* close to the truffle parasite *Tolypocladium ophioglossoides*, and *Hirsutella minnesotensis*, another nematophagous fungus. Ascomycete nematopathogenicity is polyphyletic; *D*. *coniospora* represents a branch that has not been molecularly characterized. A detailed *in silico* functional analysis, comparing *D*. *coniospora* to 11 fungal species, revealed genes and gene families potentially involved in virulence and showed it to be a highly specialized pathogen. A targeted comparison with nematophagous fungi highlighted *D*. *coniospora*-specific genes and a core set of genes associated with nematode parasitism. A comparative gene expression analysis of samples from fungal spores and mycelia, and infected *C*. *elegans*, gave a molecular view of the different stages of the *D*. *coniospora* lifecycle. Transformation of *D*. *coniospora* allowed targeted gene knock-out and the production of fungus that expresses fluorescent reporter genes. It also permitted the initial characterisation of a potential fungal counter-defensive strategy, involving interference

nih.gov/nuccore/JYHR00000000). The version described in this paper is version JYHR01000000.

Funding: This work was supported by a program grant from the ANR (ANR-12-BSV3-0001-01) and institutional funding from INSERM, CNRS and AMU. The Functional Genomics Platforms of Nice Sophia Antipolis and Toulouse are supported by the Program "Investissements d'Avenir" managed by the National Research Agency (ANR-10-09-INBS) with support from the National Infrastructure France Génomique (ANR-10-INBS-09-03 and ANR-10-INBS-09-02), the Cancéropôle PACA, the Fondation pour la Recherche Médicale (DEQ20130326464 to PB) and the Labex Signalife (ANR-11-LABX-0028-01); microscopy used the France-BioImaging infrastructure supported by ANR-10-INSB-04-01; worm sorting infrastructure supported by ANR-11-LABX-0054 (Investissements d'Avenir–Labex INFORM) and ANR-11-IDEX-0001-02 (Investissements d'Avenir–A*MIDEX). SR is supported by a Marie Curie grant (MC-CIG 334036 project SEPAraTE), a starting grant of the European Research Council (ERC-StG 336808 project VariWhim) and the French Laboratory of Excellence project TULIP (ANR-10-LABX-41; ANR-11-IDEX-0002-02). LDH received a China Scholarship Council fellowship, and NT supported by an AMU doctoral fellowship and Labex INFORM. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

with a host antimicrobial mechanism. This high-quality annotated genome for *D. coniospora* gives insights into the evolution and virulence of nematode-destroying fungi. Coupled with genetic transformation, it opens the way for molecular dissection of *D. coniospora* physiology, and will allow both sides of the interaction between *D. coniospora* and *C. elegans*, as well as the evolutionary arms race that exists between pathogen and host, to be studied.

Author Summary

Some soil-living fungi can kill nematodes and are used as biocontrol agents against plant parasitic nematodes. Certain species trap their prey using adhesive knobs or nets. For others, like *Drechmeria coniospora*, infection starts with the adhesion of specialized non-motile spores to the nematode cuticle. We have sequenced and annotated the *D. coniospora* genome. Comparative and functional genomic analyses provide insights into how its nematode-destroying lifestyle has evolved. We identified genes that were found only in *D. coniospora*, others found only in nematophagous species; many were highly expressed and differentially regulated during the different stages of fungal growth or during nematode infection. We have also developed methods for the genetic modification of *D. coniospora* that can be used to probe the function of its genes, allowing the dissection of this mode of nematode killing. We used them to probe a specific interaction between *D. coniospora* and *C. elegans*, involving the potential interference by the pathogen of a host antimicrobial mechanism.

Introduction

Species of nematophagous fungi have evolved a variety of strategies to invade and kill their nematode hosts in order to use them as a source of nutrients. Some species, exemplified by *Arthrobotrys oligospora*, form specialized and elaborate hyphal structures that trap nematodes, while others, like *Monacrosporium haptotylum*, use adhesive branches [1]. In addition to the fundamental interest in understanding these remarkable adaptations, these fungi are of economic importance because of their long-recognized potential as biocontrol agents of plant parasitic nematodes [2]. Insights into the molecular mechanisms that underlie the virulence of nematophagous species, and into their evolution, have been obtained from a series of genomic analyses (e.g. [3–9]).

Drechmeria coniospora produces non-motile spores (conidia) that stick to the nematode cuticle, via a specialized adhesive bud [10, 11]. Shortly after, the spores germinate, producing an appresorium that allows the fungus to pierce the nematode cuticle and send hyphae into its epidermis [10]. Until now, there has been essentially no molecular characterisation of *D. coniospora*. Thus, at the start of this project, nothing was known about its genetic makeup, apart from 1.05 kb of rRNA sequence in Genbank (GI:16763389; AF106012) that had been used to assign *D. coniospora* to the hypocrealean family, Clavicipitaceae, which includes many fungal pathogens of arthropods, such as *Beauveria bassiana* [12]. The same single sequence was used in a subsequent analysis that removed *Drechmeria* from the Clavicipitaceae and recognized it as one of six genera within the Ophiocordycipitaceae [13].

D. coniospora was adopted as a model fungal pathogen of *C. elegans* 30 years ago. Since the first studies in this domain [14, 15], *C. elegans* has emerged as a powerful model system for the investigation of host-pathogen interactions [16–22], and *D. coniospora* shown to be a natural

pathogen of *C. elegans* [23]. We have put considerable effort into understanding the host defences that are triggered by *D coniospora* infection (e.g. [24–28]). Great strides in dissecting host defences in other organisms have been gained by investigating how pathogens evade or subvert these mechanisms (e.g. [29–32]). Understanding what is happening on the pathogen side during infection in the *D. coniospora-C. elegans* model could therefore be key to unravelling completely the host defence network, especially as the two protagonists are likely to have co-evolved [23, 33].

Completing a high-quality draft genome of *D. coniospora* is a very useful first step for understanding its virulence mechanisms. Combined with RNAseq transcriptomic and *in silico* analyses, it allowed us to predict a first complete gene set for *D. coniospora*. A comparison with other fungi, including nematode-destroying species, has revealed genes potentially involved in virulence and given insights into the evolution of the infectious capability of *D. coniospora*. To be able to exploit this knowledge, we established a method for fungal transformation, and, in a proof-of-principle, used it to generate recombinant knock-out and knock-in strains. These different approaches allowed us to initiate the investigation of a potentially novel fungal counter-defensive strategy.

Results

Sequencing output processing, *de novo* genome assembly and scaffolding

D. coniospora genomic DNA was sequenced on an Illumina MIseq sequencer as 2 x 150 bp paired-end reads. After filtering, we obtained 11.3 million reads, for a 100X coverage of a genome originally estimated to be 30 Mb. To determine which frequently used *de novo* genome assembly program performed best with this set of data, we tested four, Velvet [34], SPAdes [35], SOAPdenovo2 [36] and ABySS [37]. Each assembly was scaffolded with SSPACE [38], using two libraries of mate-paired 2 x 60 bp SOLiD reads. We applied a stringent filter, keeping only very high-quality reads, to limit errors during scaffolding. The libraries finally contained 23.2 and 23.6 million mate-paired reads, with insert sizes of 1.5 kb and 3 kb, respectively. Two contigs were scaffolded, using SSPACE, only when supported by at least 5 shared mate-paired reads. The overall characteristics of the *de novo* genome assemblies are shown in Table 1, before and after the SSPACE scaffolding step. ABySS (with kmer 96) and SPAdes performed well, giving low numbers of both contigs and unknown nucleotides, with ABySS maximizing the N50 value (2.09 Mb; length for which the collection of contigs of that length or longer contains half the total length of all contigs).

Optical mapping data integration

We then used optical mapping to test further the quality of the different assemblies generated before and after SSPACE scaffolding. Individual chromosomes were stretched on a glass slide and cut *in situ* with a restriction enzyme. The resulting fragments were visualized using a fluorescent microscope and their lengths measured. These lengths were compared to the predicted lengths of fragments from the longer scaffolds (i.e. with a length of at least 20 kb). The same approach was applied for each assembly. The nine distinct maps that were identified by optical mapping are indicative of a genome organization into 9 distinct chromosomes, with sizes ranging from 0.58 to 11.3 Mb (S1 Table).

The assembly obtained with ABySS (kmer = 96) was selected since it maximized the remapping of the scaffolds on the optical map whilst at the same time minimizing the number of misassemblies observed in the optical map analysis (4 versus 20 for Spades and 9 for ABySS at a

Table 1. Descriptive statistics of the different assemblies before and after SSPACE scaffolding.

	Before SSPACE					After SSPACE				
	Vel	So	Sp	A64	A96	Vel	So	Sp	A64	A96
Number contigs	455	12451	1075	7880	1356	190	11700	470	7157	1023
Total length (Mb)	31.6	33.1	31.9	32.3	32.0	31.7	34.0	31.9	32.8	32.1
Max contig length (Mb)	1.27	1.94	0.51	1.36	1.20	3.00	4.54	4.50	2.83	4.80
N50 (Mb)	0.32	0.58	0.12	0.35	0.24	1.40	1.76	1.21	1.70	2.09
% of Ns	0.77	1.85	0.00	0.24	0.12	1.13	4.58	0.16	1.85	0.33

Total number of contigs; total size of the assembly; length of the longest contig; N50; percentage of unknown bases. Values are given for each assembler (Vel: Velvet with kmer = 63, So: SOAPdenovo with kmer = 63, Sp: Spades with kmer = 127; A64 and A96: ABySS with kmer = 64 and 96, respectively) before and after scaffolding with SSPACE.

doi:10.1371/journal.pgen.1006017.t001

kmer value of 64; S1 Table). This low occurrence of scaffolding errors following SSPACE indicates that the first scaffolding step based on SOLiD reads would probably already have been of sufficient quality for an accurate assembly. During optical map analysis integration, we decided not to join scaffolds located on the same optical map when separated by a gap larger than 150 kb. Rather, for omap6937 and omap49267, the resulting map was divided into sub-chromosomal sequences. For instance, omap49267, the longest map, has a length of 11.3 Mb. It is represented in the final assembly by 2 sub-chromosomal sequences: omap49267a, with a length of 1.6 Mb and omap49267b, with a length of 10 Mb. The discrepancy in size between the sum of the lengths of the 2 sub-chromosomal sequences and that of omap49267 (11.6 Mb instead of 11.3 Mb) is explained by the fact that we did not break the scaffolds' distal extremities during scaffold concatenation. It is noteworthy that the sequences both at the 3' end of omap49267a and the 5' end of omap49267b (i.e. within the corresponding chromosome) are of low complexity (Fig 1), which can explain the difficulty in mapping properly these extremities on the optical map. We then put aside all contigs smaller than 0.5 kb (a collection totalling 128 kb). The final genome assembly therefore includes 75 sequences for a total size of 31.9 Mb. It contains less than 0.2% of unknown nucleotides, and its N50 value is equal to 3.86 Mb. The combination of 2 types of short-read sequencing (SOLiD and MISeq) with optical mapping therefore provided the basis for a high-quality assembly of the *D. coniospora* genome.

Transposable elements, repeat sequences, tRNAs, rRNAs and mtDNA

Amplification of transposable elements (TE) can contribute to chromosomal rearrangements, altering genomic structure and gene expression. It is believed to be an important route to speciation in some fungi. We therefore characterized the set of TEs predicted in the *D. coniospora* genome using TransposonPSI (http://transposonpsi.sourceforge.net/). A total of 600 elements were detected, falling into 12 classes (Table 2), and covering 1.6% of the genome (516 kb).

We compared these results with the number of TEs found in the genomes of 11 other fungi, chosen on the basis of their phylogenetic position and/or lifestyle, using the name assigned by the NCBI Taxonomy Database: (1) *Arthrobotrys oligospora*, *Hirsutella minnesotensis*, *Monacrosporium haptotylum* and *Pochonia chlamydospora*, 4 nematophagous fungi that feed on different species and stages of nematode worms; (2) the entomopathogenic fungi *Metarhizium acridum*, *Metarhizium anisopliae* and *Ophiocordyceps sinensis*; (3) the plant pathogens *Fusarium graminearum* and *Fusarium oxysporum*; (4) the mycoparasite *Tolypocladium ophioglossoides*; (5) *Trichoderma reesei*, a model mesophilic and filamentous fungus. Although being numerous in *D. coniospora*, the number of TEs was in no way exceptional. For example, there

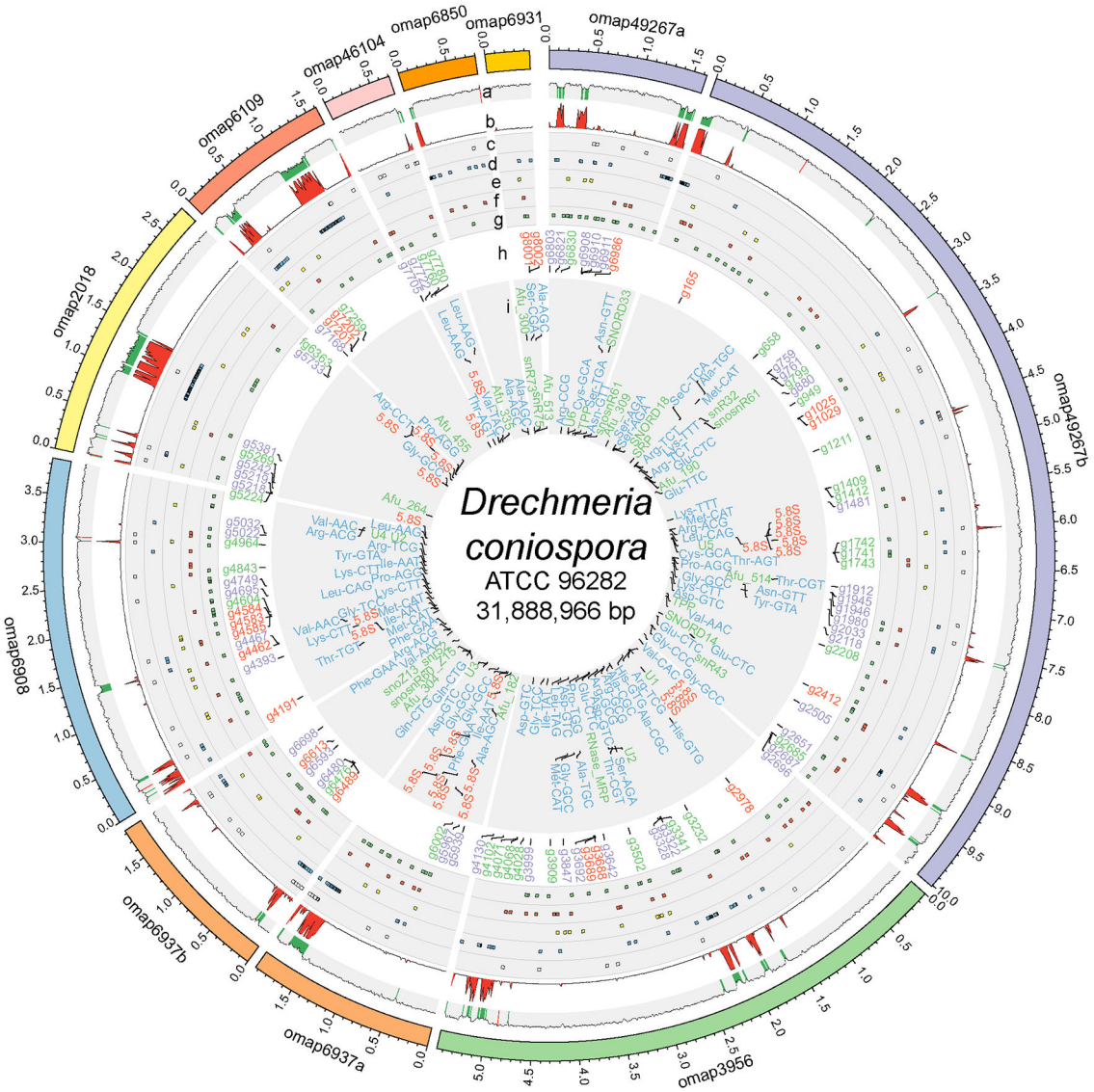


Fig 1. The *Drechmeria coniospora* genome. The optical maps (omap), potentially corresponding to distinct chromosomes, are depicted using Circos [39] with coloured sectors on the outer layer. As explained in the text, the optical maps 6937 and 49267 are split into two pieces. Scaffold43, corresponding to 23.8kb of mitochondrial DNA, and a further 63 other unanchored scaffolds, including 41 containing at least one predicted gene, and totalling 755,493 bp of genomic sequence, are not printed on the Circos plot. Each layer depicts, from the outside to the inside: **(a)** Percentage of G+C (red > 0.65, green < 0.45); **(b)** Percentage of repeat elements (red > 10%); **(c and d)** CLASS II and CLASS I transposable elements (white and blue blocks, respectively); **(e)** Members of three superfamilies encoding glutathione-S-transferases (GSTs), cytochrome P450 monooxygenases (P450s) and carboxyl/cholinesterases (CCEs) important for xenobiotic detoxification and oxidative stress resistance in entomopathogenic species [40] are depicted in yellow; **(f)** TM7 transmembrane proteins (red); **(g)** ABC proteins (green); **(h)** groups of genes discussed in the text that encode: putative nonribosomal peptide synthetases (red), diverse proteases (purple) and enterotoxin-like proteins in green (their names without the.t1 suffix are shown); **(i)** non-coding RNA genes: tRNAs (blue), rRNAs (red), others (green).

doi:10.1371/journal.pgen.1006017.g001

has been a remarkable proliferation of retrotransposons in *Ophiocordyceps sinensis* [41]; its genome is predicted to contain 12862 TEs (S2 Table).

As part of its overall characterisation, we mined the *D. coniospora* genome for repetitive elements using RepeatScout [42]. The resulting specific repeat library was used for masking the genome with RepeatMasker (www.repeatmasker.org) and finally covered 2.73 Mb (8.6%) of

Table 2. Transposable elements in the *D. coniospora* genome.

Class 1 DNA transposons, LTR retroelements		Class 2 DNA transposons			
Name	**Number**	**Name**	**Number**	**Name**	**Number**
TY1_Copia	193	hAT	38	MuDR_A_B	15
gypsy	193	cacta	41	piggybac	2
DDE_1	46	mariner	23	helitronORF	2
LINE	26	mariner_ant1	20	ISC1316	1

doi:10.1371/journal.pgen.1006017.t002

the full genome. Genes for ribosomal 28S and 18S RNAs were identified on scaffold58 (6.3 kb), near a 5.8S rRNA subunit gene. In addition, 26 copies of 5S RNAs were detected scattered throughout the entire genome. A further 33 non-coding, non-tRNAs/rRNAs were detected. We also searched for predicted tRNAs and found a total of 123 nuclear tRNA loci (Fig 1). Eighteen mitochondrial-specific tRNAs were located on Scaffold43 (23.8 kb, G+C 25.7%) after a tRNAscan search [43] with the option "organelle". The 141 predicted tRNAs correspond to 49 out of 64 codons. This degree of coverage compares very favourably with other sequenced fungal genomes (S3 Table). A more comprehensive investigation of Scaffold43 with TBLASTX [44] against a database of mitochondrial genes led to the identification of homologues for multiple mitochondrial proteins and established it as the mitochondrial DNA (S1 Text).

Gene prediction and genome annotation

From the genomic DNA, Augustus [45] was used to predict the coding DNA sequence (CDS) for a total of 8733 genes, with a mean length of 1425 bp and a GC content of 61.3% (higher than for the overall genome sequence, as is generally the case). The predicted set represents a total of 12.4 Mb of coding nucleotides. To assess the quality and completeness of the prediction, we performed a BUSCO analysis [46], which is based on expectations of gene content from near-universal single-copy orthologs (USCOs). Using a set of more than 1400 fungal USCOs, again the annotation of the *D. coniospora* genome appeared to be of high quality, at least as good as that of the other fungi used in this study (Table 3).

With these results in mind, it is therefore interesting to compare this first comprehensive prediction for the *D. coniospora* genome with those of the other 11 fungi (Table 4).

D. coniospora was recently reassigned to the family Ophiocordycipitaceae [13]. Overall, compared to the 3 other family members included in our analysis, the *D. coniospora* genome much more closely resembles that of *Tolypocladium ophioglossoides* than either *Hirsutella minnesotensis* or *Ophiocordyceps sinensis*, in terms of size, GC content and the number of predicted genes. Both species have comparatively small genomes but a relatively high complement of predicted protein-coding genes.

Phylogeny

A previous study placed *D. coniospora* in Ophiocordycipitaceae on the basis of a single DNA sequence [13]. In order to carry out a more thorough phylogenetic analysis, using BUSCO, we identified a set of 97 high-confidence orthologous proteins present in all 12 fungal species. Concatenated sequences (S4 Table) were aligned using MAFFT [47] and phylogenetic distances calculated using PhyML [48]. The overall phylogeny was in line with recent phylogenetic studies of *H. minnesotensis* and *P. chlamydosporia* [9, 49], and confirmed *D. coniospora*'s place in the Ophiocordycipitaceae family. Consistent with the general features of their respective genomes, the analysis placed *D. coniospora* closest to *T. ophioglossoides* (Fig 2). The results

Table 3. Universal single-copy ortholog prediction in *D. coniospora* and 11 other species.

	Complete single-copy	Complete duplicated	Fragmented	Missing	Source[a]
Drechmeria coniospora	98% (1412)	10% (153)	1.5% (22)	0.2% (4)	This study; PRJNA269584
Trichoderma reesei	97% (1407)	11% (167)	1.6% (24)	0.4% (7)	Trire2/Trire2.home.html
Fusarium graminearum	97% (1406)	11% (167)	2% (29)	0.2% (3)	Fusgr1/Fusgr1.home.html
Arthrobotrys oligospora	97% (1404)	11% (159)	2.1% (31)	0.2% (3)	Artol1/Artol.home.html
Monacrosporium haptotylum	96% (1390)	10% (156)	2.7% (39)	0.6% (9)	Monha1/Monha1.home.html
Hirsutella minnesotensis	96% (1388)	12% (183)	2.3% (34)	1.1% (16)	PRJNA67943
Fusarium oxysporum	96% (1386)	35% (511)	3.2% (47)	0.3% (5)	Fusox1/Fusox1.home.html
Metarhizium anisopliae	95% (1376)	11% (165)	3.5% (51)	0.7% (11)	Metan1/Metan1.home.html
Tolypocladiumophioglossoides	95% (1374)	10% (157)	1.6% (24)	2.7% (40)	PRJNA91059
Metarhizium acridum	95% (1370)	10% (152)	4% (58)	0.6% (10)	Metac1/Metac1.home.html
Pochonia chlamydosporia	88% (1270)	11% (168)	9% (130)	2.6% (38)	www.fungalinteractions.org/index.php/en/genome
Ophiocordyceps sinensis	71% (1026)	7.3% (105)	10% (152)	18% (260)	PRJNA59569

The table shows the percentage of the different categories of USCOs, with the corresponding number of proteins in brackets, as calculated by BUSCO. Orthologs are classified as 'complete' when their lengths are within two standard deviations of the BUSCO group mean length, otherwise they are classified as 'fragmented' (length not within the threshold) or 'missing'. 'Complete' orthologs found with more than one copy are classified as 'duplicated'.
[a]NCBI bioproject number, full URL, or end of URL at genome.jgi.doe.gov

doi:10.1371/journal.pgen.1006017.t003

support the conjecture that in Hypocreales invertebrate-pathogenic fungi form a monophyletic group, distinct from cellulolytic, plant pathogenic filamentous fungi [49], while also providing a further illustration of the multiple independent origins of nematode pathogenic fungi and the distinct evolutionary trajectories of the trapping fungi such as *Arthrobotrys oligospora*, as opposed to the conidial species including *D. coniospora*.

Functional annotation of the predicted proteome

In order to obtain a first overview of the set of proteins predicted from the *D. coniospora* genome, we functionally annotated the protein sequences with InterproScan [50]. This assigned at least one annotation to more than three quarters of them (6734/8733; 77.1%). There was a bias in the distribution of annotations; longer proteins were more likely to have an annotation. Thus while the vast majority (90.7%) of proteins as long or longer than the median (393 amino acids) had an annotation, only 63.4% of proteins shorter than the median had one. For smaller proteins, the effect was even more marked. Indeed, only half (50.6%) of proteins shorter than 250 amino acids long had an annotation (S5 Table). In line with previous observations [51], this bias was mirrored in the pattern of conservation. While overall, 60% (5248/8733) of the predicted proteins have a homologue in the curated UniprotKB/Swissprot database, 74.8% of the proteins as long or longer than the median were assigned a homologue, while only 34.6% of the proteins less than 250 amino acids were. It should be noted that these are conservative estimates of homology since although of high quality, the UniprotKB/Swissprot database is not exhaustive. Indeed, for example, on the basis of BLASTP analyses using the current NCBI non-redundant database, many (19/31) of the predicted proteins longer than 1000 amino acids but without any annotation have homologues in other fungi. The closest homologue was most often found in *Tolypocladium ophioglossoides* (S5 Table), consistent with the phylogenetic analysis.

Table 4. General characteristics of the *D. coniospora* genome compared to other species.

	Drechmeria coniospora	*Metarhizium acridum*	*Metarhizium anisopliae*	*Pochonia chlamydosporia*	*Arthrobotrys oligospora*	*Monacrosporium haptotylum*	*Fusarium graminearum*	*Fusarium oxysporum*	*Trichoderma reesei*	*Tolypocladium ophioglossoides*	*Hirsutella minnesotensis*	*Ophiocordyceps sinensis*
Sequencing depth (fold coverage)	100	107	100	136	37	28	85	-	48	76	128	241
Genome size (Mb)	31.9	39.4	39.2	42.4	40.1	39.5	36.5	61.4	33.5	31.2	51.1	78.5
Chromosomes[a]	9 + MT	-	-	-	-	-	4 + MT	15 + MT	7 + MT	-	-	-
GC content (%)	55.2	49.9	51.5	49.9	44.5	45.3	48.3	48.4	52.8	57.3	52.1	43.5
# of scaffolds	75	241	176	956	215	1279	31	114	87	172	736	10603
Scaffold N50 (Mb)	3.86	0.33	1.96	0.23	2.04	0.19	5.35	1.98	1.22	0.67	0.38	0.01
% of unknown (N's)	0.2	3.49	0.27	2.91	0.27	0.09	0.61	2.32	0.14	3.15	2.44	5.77
No. predicted protein-coding genes	8733	9849	10583	11079	11479	10959	13322	17708	9849	9317	12700	6972
Average exons per gene	2.83	2.7	2.68	-	3.17	3.31	2.82	2.7	2.88	-	2.5	-
Average exon length (bp)	504	549	568	-	473	470	508	498	512	-	-	-
Median protein length (aa)	393	406	416	393	407	416	366	366	408	410	403	362
Mitochondrial DNA (kb) [a]	23.8	145	24.7	25.6	170.4	140.2	107.7	84.8	42.1	-	40.6	42.2
Repetitive sequence in Mb (%)	2.73 (8.6)	1.16 (2.9)	0.59 (1.5)	0.06 (0.1)	0.54 (1.4)	0.73 (1.9)	0.35 (1.0)	11.85 (19.3)	0.17 (0.5)	0.37 (1.2)	13.22 (25.9)	51.64 (65.8)
Transposable elements in Mb (%)	0.52 (1.6)	0.15 (0.4)	0.33 (0.8)	0.14 (0.3)	0.17 (0.4)	0.32 (0.8)	0.04 (0.1)	3.176 (5.2)	0.11 (0.3)	0.07 (0.2)	3.64 (7.1)	8.51 (10.8)
NCBI accession	PRJNA269584	ADNI00000000.1	PRJNA156697	AOSW00000000	ADOT00000000.1	AQGS00000000.1	AACM00000000.2	AAXH01000000	PRJNA118357	PRJNA91059	PRJNA67943	PRJNA59569

[a] See S1 Text

doi:10.1371/journal.pgen.1006017.t004

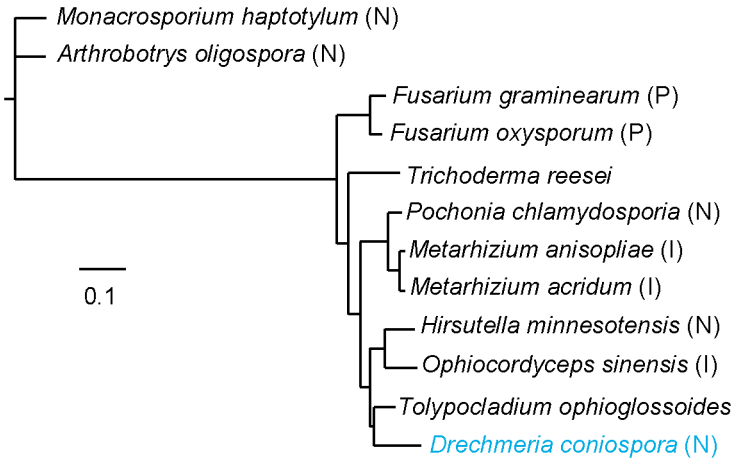


Fig 2. Phylogenetic tree for 12 Pezizomycotina fungi. Phylogenetic tree for 12 species based on alignments for a concatenation of 97 conserved protein sequences. Branch-lengths are drawn in proportion to the estimated number of substitutions per site. Species known to infect insects (I), nematodes (N) and plants (P) are indicated. All branches are fully supported (100/100 bootstraps).

doi:10.1371/journal.pgen.1006017.g002

To define gene families and assign proteins to orthologous and paralogous groups, we performed an OrthoMCL analysis [52]. Most proteins (6851) were assigned to an OrthoMCL orthologous group, together with 241 for which an orthologue but no group was assigned (in total 81.2%; S6 Table). For the great majority (82%) of proteins allocated to an OrthoMCL orthologous group, the closest homolog was in *Fusarium graminearum* (the anamorph name of *Gibberella zeae*), the only Hypocrealean fungal species represented in OrthoMCL (S6 Table). The most populated groups, with 20 members (OrthoMCL group OG5_126718) correspond to predicted nonribosomal peptide synthases (see below), followed by ABC transporters and subtilisin-like serine proteases (OG5_134254 and OG5_137388, respectively, with 15 members each). When no homolog is found in any of the 150 species in OrthoMCL, proteins are placed into paralogous groups on the basis of their sequence [52]. The largest group of paralogs, with 15 members, corresponds to the heat-labile enterotoxin alpha chain, present in several insect pathogenic fungal species including *B. bassiana* and *Metarhizium robertsii* (e.g. EXU96489) [53]. A further 3 groups (containing a total of 18 predicted proteins) exhibited a more-or-less strong similarity to the heat-labile enterotoxin alpha chain, and one group of 5 to subtilisin-like serine proteases, also mentioned above (Tables 5 and S6). As discussed below, these all potentially play a role in fungal virulence or in the interaction of *D. coniospora* with other microbes. This analysis also revealed many groups of proteins currently unique to *D.*

Table 5. Families of predicted *D. coniospora* paralogs.

OrthoMCL group(s)*	Number of members	Identity
1, 4, 9, 17	15, 8, 6, 4	Heat-labile enterotoxin alpha chain
2, 5, 6, 7, 8, 10, 14, 15, 16, 19	10, 8, 7, 7, 6, 6, 4, 4, 4, 4	*Drechmeria*-specific; unknown function
3, 12, 13, 18, 20	9, 5, 4, 4, 4	Hypothetical protein conserved in certain fungal species; unknown function
11	5	Subtilisin-like serine protease

*See S6 Table

doi:10.1371/journal.pgen.1006017.t005

coniospora. In the most dramatic example, a group of proteins (OrthoMCL paralogous group 2) with no recognisable domain, each entirely composed of highly repeated sequences, has expanded to 10 members, with the majority found in a cluster of less than 50 kb on scaffold omap6908, with a very complex pattern of conservation (S6 Table and Figs 3 and S1). A further 9 out of the 20 groups with at least 4 members currently correspond to proteins unique to *D. coniospora* (Table 5).

Comparative analysis of protein domains

To look in more detail at protein domains, we next compared the PFAM annotations [56] for the entire set of predicted *D. coniospora* proteins with those of the 11 other fungi obtained with InterproScan using the same set of parameters (S7 Table). We completed the analyses of specific proteins and protein families by manual inspection. A total of 4287 PFAM domains were identified in at least one protein from one or more of the species, with 3510 (81.9%) represented in predicted *D. coniospora* proteins and more than half (55.7%) in all 12 predicted proteomes. These presumably reflect core eukaryotic and/or fungal biological processes. In the expectation of revealing domains that were potentially functionally related, we hierarchically clustered protein domain families present in *D. coniospora* and not more than 4 of the other fungi, but no obvious associations were found (Fig 4). A number of PFAM domains were predicted for *D. coniospora* proteins but not for proteins of any of the other 11 fungi (Fig 4 and S7 Table and S1 Text). Within this group, each PFAM domain is present in a single protein, with 3 exceptions. Three predicted *D. coniospora* proteins (OrthoMCL paralogous group 33, S6 Table and S1 Text) contain PF12810, the "glycine rich protein" domain, characterised by several glycine rich motifs interspersed through the sequence. Currently, no orthologues have been described in any other species, and no hint of a function can be garnered from the sequence. Two lipid-binding MORN (Membrane Occupation and Recognition Nexus; PF02493) domains [57] are predicted, towards the C-terminus, in proteins from each of 2 adjacent *D. coniospora* genes (g5037.t1 and g5038.t1; OrthoMCL orthologous group OG5_154358). MORN domains are relatively uncommon in fungi, but tandemly arranged orthologues for these 2 proteins do exist in one species, *Trichoderma gamsii*, and orthologs are currently also found in various other fungi, such as the brown-rot Basidiomycota *Hydnomerulius pinastri*. The conserved N-terminal portion of these proteins is shared with a number of related toxins, including the hemolytic factor neoverrucotoxin from stonefish venom (S5 Table). Whether these proteins may play a role in fungal virulence and interactions with other microorganisms is a matter for speculation. The presence of the Saposin A domain (PF02199) in 2 predicted proteins (g3895.t1 and g1982.t1, with 3 and 2 domains respectively) is equally atypical (S5 and S7 Tables). We address its possible role below.

On the basis of their constituent domains, several other atypical or highly represented protein families (Fig 4 and S7 Table) are also potentially linked to virulence. These include the single iron-sequestrating lipocalin (PF13924) and the deuterolysin M35 metalloprotease (PF02102; called here M35) domains. The M35 domain is unusually highly represented compared to other fungi [58, 59], being present in 10 predicted *D. coniospora* proteins.

Of a total of 777 domains absent from *D. coniospora*, 32 were present in all 11 other species (S7 Table). As a most striking example, *D. coniospora* lacks proteins containing the NACHT domain (PF05729), which is present between 8 and 117 times in the other species, suggesting that this is unlikely to simply be a problem of sequence coverage or gene prediction. In ascomycete fungi, the NACHT domain can be found together with the HET domain (PF06985) in heterokaryon incompatibility proteins. It acts as a death effector domain, preventing viable heterokaryotic cells from being formed by the fusion of filaments from different wild-type strains

A

```
5-41      LRNVHYCNNAAY.FERYDN....NNHRTHCRYINDAANPWMY
44-80     LHNFYFCYNASS.LKRNNI....YNSGPNCRYFDCLAASWLY
83-119    LHNDHYHHNASYiWGNNNIharpNRR.....YIDDFAISRMC
122-158   LLYDHYCYDNSY.FESNNK....DNTWPNCRYVDDPPACWMY
161-197   LCNVHYHHDTSY.LGRNND....IHTRANRRFIDDPANTRMY
200-236   LLNDHYCHDTSY.FGRNNI....NDTWSYCWYVDNLAACWLY
239-275   LHDVNYCYNTSY.VKSNNK....DNTRPNCRYVDYPTASWMY
278-314   LRNIHYHHDTSY.LGRNND....IHTRANRRFIDNPANTRMY
317-353   LPNDHYCHDTSY.LGRNNI....NDTWSYCWYVDNLAACWLY
356-392   LHDVNYCYNTSY.VKSNNK....DNTRPNCRYVDYPTASWMY
395-431   LRNVHYHYDTSC.LERNNI....NYTRSNCWHIYDLAARWLY
434-470   LYNVHYCHNASR.FQWNNI....DDTRSDCRYIDNLATSRMY
```

B

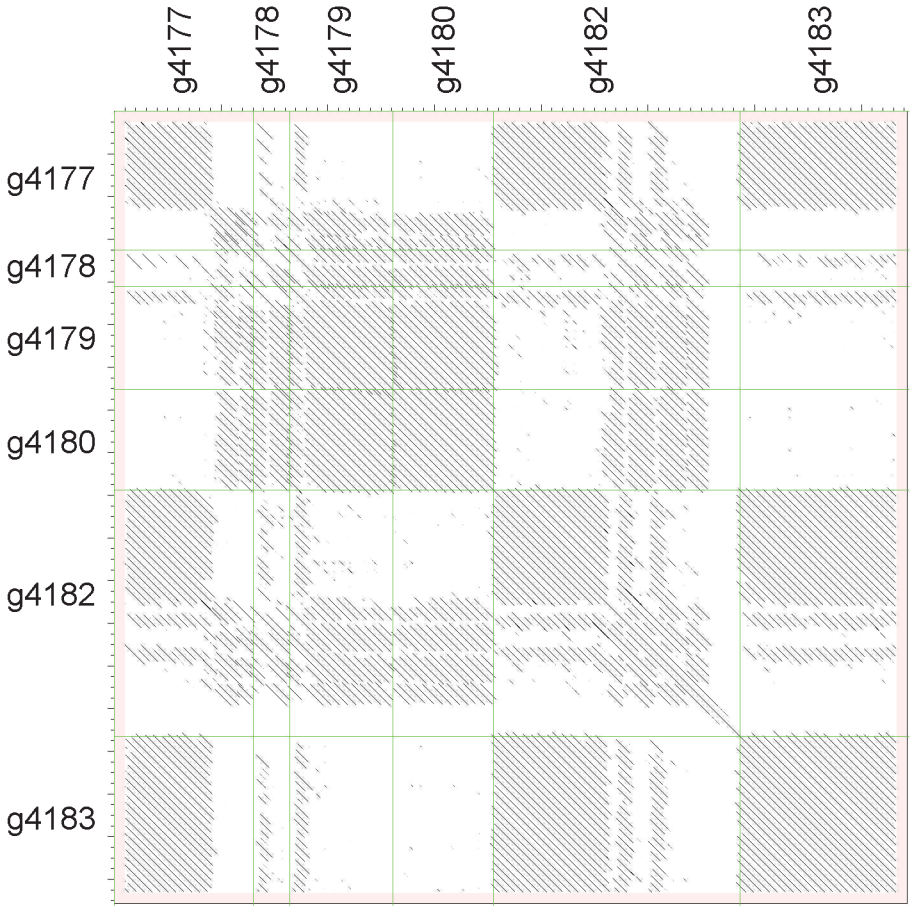


Fig 3. Unusual structure of *Drechmeria*-specific proteins and complex structural relationship between neighbouring proteins. (A) RADAR analysis [54] reveals the repeated structure in the sequence of g4180.t1, a 471 a.a. protein from OrthoMCL-defined paralogous group 2 (S6 Table). (B) All-against-all dot-plot representation [55] of the alignment of the predicted protein sequences from g4180.t1 and from 5 neighbouring genes on scaffold omap6908, all from the OrthoMCL paralogous group 2. Dots represent regions of sequence similarity (within a 100 a.a. sliding window). The intensity of each dot is proportional to the corresponding alignment score. The ".t1" suffix has been removed from all sequence names.

doi:10.1371/journal.pgen.1006017.g003

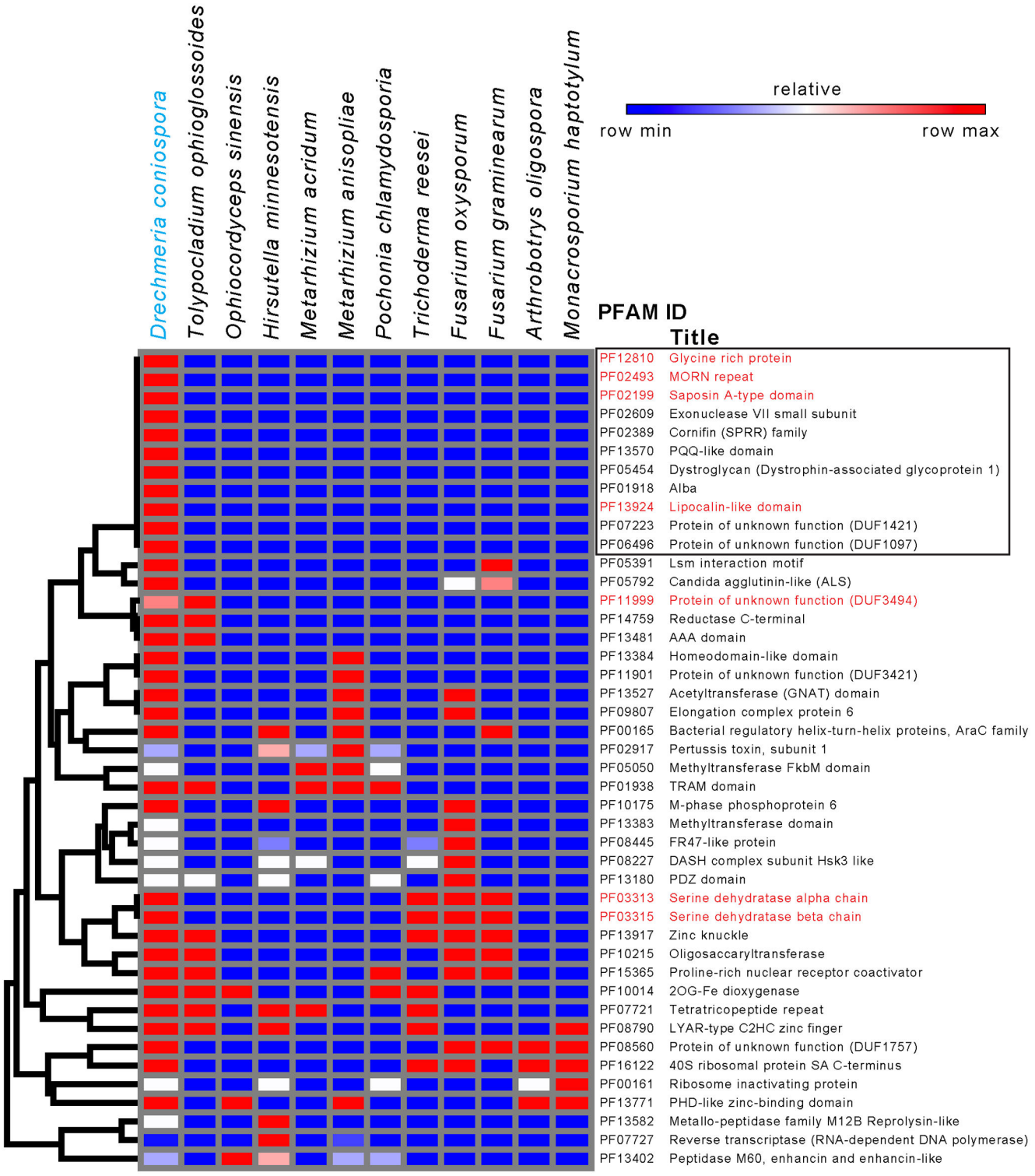


Fig 4. Species-specific or atypical protein domain families in the predicted *D. coniospora* proteome. Hierarchical clustering of protein domain families present in *D. coniospora* and not more than 4 of the other fungi, on the basis of the corresponding number of proteins. PFAM domains discussed in the text are highlighted in red. The box highlights 11 families specific to *D. coniospora* (see S1 Text). The colour code reflects the relative abundance of proteins with each domain, from high (red) to low (blue) across the different species. The serine dehydratase alpha and beta domains (PF03313 and PF03315, respectively), cluster since they occur in a single highly conserved protein (g4699.t1 in *D. coniospora*).

doi:10.1371/journal.pgen.1006017.g004

[60]. In *D. coniospora*, of the 6 proteins predicted to contain a HET domain, 5 (g3856.t1, g5845.t1, g5969.t1, g6809.t1, g7038.t1) do not contain any other domains, while in common with many other fungal proteins of this class, one (g215.t1) is coupled to an ankyrin repeat (PF12796) domain. No sexual stage has been described for *D. coniospora*, and in common with 3 of the 4 other nematophagous fungi, no mating type protein (MATα, PF04769) domain was predicted (S7 Table), making it unlikely that these different proteins play any role in heterokaryon incompatibility.

To see whether there was any general pattern for the 32 domains absent from *D. coniospora* suggestive of a coordinate evolutionary process, we used dcGO analysis to look for ontology term enrichment at the domain level [61]. This revealed a potential involvement of the NACHT domain together with 4 others (PF00931; PF01966; PF02178; PF09273) in the biological process, "regulation of response to stress" and of 4 of them (PF00931; PF01966; PF05729; PF09273) in "regulation of defense response". This is discussed further below. Submitting all 777 domains absent from *D. coniospora* to dcGO analysis highlighted the absence of 16 related domains, all annotated as being involved in hydrolase activity, acting on glycosyl bonds, including PF00331 that corresponds to the Carbohydrate-Active Enzyme (CAZyme) glycosyl hydrolase (GH) family 10. There was also a significant ($p = 2.5x10^{-4}$) enrichment for the more specific ontology term "alpha-N-arabinofuranosidase activity" (PF05270, PF06964, PF09206), suggesting an alteration in the capacity of *D. coniospora* to metabolise different carbohydrates compared to other fungi analysed.

Carbohydrate-active enzymes

A species' set of CAZymes can often give insights into its biology, in particular into nutrient sensing and acquisition. Given the differences revealed by the dcGO analysis, we conducted a targeted examination of CAZymes in *D. coniospora* and 10 of the 11 fungi chosen for the other comparative analyses (S8 Table). *P. chlamydiosporia* was not included as it will be the subject of a dedicated study. Overall, CAZyme profiling recapitulated the phylogenetic analysis, except that the two *Metarhizium* species clustered together with *H. minnesotensis* and *O. sinensis* (Fig 5). This grouping of the various fungi reflects their respective requirements for carbon acquisition. The two nematode-trapping fungi *A. oligospora* and *M. haptotylum* have a large repertoire of enzymes for feeding on plant cell wall polysaccharides. They make a separate group and are neighbours of the saprophytes, reflecting their dual parasitic and saprophytic lifestyles. *T. reesei* has an intermediate position consistent with its evolving from an ancestral saprophyte lifestyle to become a mycoparasite. The remaining fungi, which are the most specialized and have evolved by gene loss, group together by virtue of their common loss of an arsenal of plant polysaccharide degradative enzymes. Thus the nematophagous and insectivorous fungi in this group are not separate from the mycoparasite *T. ophioglossoides*; the same range of CAZymes is probably needed for the three types of substrate and this is accompanied by a similar loss of the plant-targeting CAZymes. Regarding *D. coniospora* in detail, it has lost virtually all enzymes, from multiple families, that participate in cellulose binding (Carbohydrate Binding Module, CBM1), the breakdown of cellulose/hemicellulose and pectin-rich plant cell walls (e.g. GH7, GH45, PL8, and CE8 family proteins; S8 Table). The few GH5 proteins that remain in *D. coniospora* are predicted to be involved in the metabolism of fungal cell wall ß-glucans, not the digestion of plant cellulose or mannan. The GH13 family, involved in both starch and glycogen breakdown, has also shrunk to just two members. The two remaining proteins show strong similarities to glycogen branching and debranching enzymes and are thus most likely involved in the fungal glycogen cycle. Collectively, and coupled with the expansion of protease families

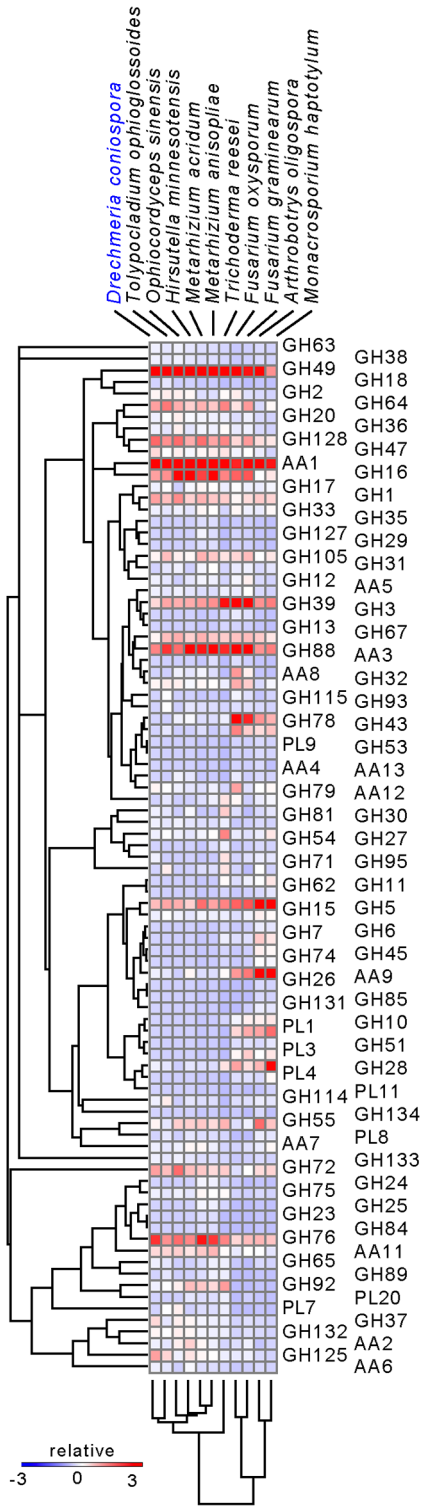


Fig 5. Supervised clustering of selected CAZy families in 11 fungal species. The distribution of the CAZy families involved in complex carbohydrate breakdown (AA, GH and PL classes) across the given species is shown. Clustering of families is based on the number of genes in each family. The colour code reflects the relative abundance of proteins within each family, from high (red) to low (blue) across an individual species.

doi:10.1371/journal.pgen.1006017.g005

and the acquisition of virulence factors, these changes appear to reflect the shift of *D. coniospora* to an obligate nematophagous lifestyle.

Secondary metabolite biosynthesis

Another hallmark of each fungal species is its complement of genes involved in the production of secondary metabolites. The "backbone" genes for these biosynthetic pathways include those encoding nonribosomal peptide synthases (NRPSs), polyketide synthases (PKSs), and prenyltransferases (DMATSs), responsible for the production of bioactive peptides, polyketides, and indole alkaloids, respectively. SMURF analysis [62] revealed that the number and type of backbone genes predicted for *D. coniospora* was comparable to those of other Hypocreales fungi (Table 6). The majority of the NRPS and NRPS-like genes belong to OrthoMCL group OG5_126718. Several of the genes (g1025.t1 and g1029.t1; g3636.t1 and g3638.t1; g7201.t1, g7202.t1 and g7204.t1; g8001.t1 and g8002.t1) are close together in the genome suggesting that they may be functionally related (S9 Table). The molecules synthesized by NRPSs, PKSs and DMATS are frequently modified by "decorating" enzymes before secretion. The genes encoding the proteins necessary for these different steps are often found in genomic clusters, and are co-ordinately regulated by specific Zn_2Cys_6 transcription factors and/or by the global secondary metabolism regulator LaeA [62]. Of the 29 backbone genes, 24 had an associated gene cluster, and of these, 4 included a Zn_2Cys_6 gene. These are therefore candidate regulators of their respective clusters. Two clusters (14 and 15) included genes in the proximity of the *D. coniospora* LaeA-encoding gene g6733.t1 (S9 Table). In *D. coniospora*, LaeA could play a conserved role in secondary metabolism. It is important to note that in many fungal pathogens, secondary metabolites are essential for virulence [63].

PHI-base analysis

To gain a general view of proteins potentially involved in virulence, we made use of the pathogen-host interactions database, PHI-base (phi-base.org; [65]). Of the 2104 proteins matching a PHI-base entry, 990 had the annotation "reduced virulence" or "loss of pathogenicity", indicating that a homologous protein in at least one other species plays a demonstrated role as a virulence factor in a particular model of infection (S5 Table).

Among the highly represented ($\geq$5) hits in PHI-base, several were characterised by the presence of ABC transporter domains, and so predicted to be involved in ATP-dependent export of organic anions or drugs from the cytoplasm (S10 Table and Fig 1). For predicted NRPS proteins, 9/20 members of the orthoMCL group OG5_126718 were assigned a PHI-base annotation (PHI:2511). The analysis also highlighted the potential role of multiple degradative enzymes, 11 chitinases (GH18; PHI:144; 6/6 OG5_126929, 2/2 OG5_142806, 2/2 OG5_210539, 1/1 OG5_152762), known to be important for the virulence of nematophagous fungi [7, 66], and

Table 6. Number of secondary metabolite backbone genes predicted from the *D. coniospora* genome compared to other species.

	DMAT	NRPS	NRPS-Like	PKS	PKS-Like	Reference
Drechmeria coniospora	1	10	9	7	2	This study
Tolypocladium ophioglossoides	0	14	6	16	2	[64]
Trichoderma reesei	0	8	5	11	1	[62]
Hirsutella minnesotensis	6	21	21	27	4	This study
Fusarium oxysporum	2	7	12	9	2	[62]
Fusarium graminearum	0	10	11	14	1	[62]

doi:10.1371/journal.pgen.1006017.t006

subtilisin-like and extracellular metalloproteases (PHI:2117 and 479, with 9 and 5 members, respectively). These are often found in expanded gene families in pathogenic fungi (e.g. [67]). There were also 11 Pth11-like receptors (PHI:404), which can be involved in host sensing and have established roles in virulence in other species [68]. Three of them contain a CFEM domain (PF05730), found in proteins with proposed roles in fungal pathogenesis [69]. As expected from the OrthoMCL analysis, there were also multiple hits to enterotoxin A proteins (PHI:698, with the PF01375 domain), scattered throughout the genome (g496.t1, g964.t1, g2819.t1, g5058.t1, g6833.t1, g7169.t1, g7949.t1). The *D. coniospora* genome is therefore predicted to encode a large range of virulence factors, some of which have expanded markedly in number compared to other fungal species.

D. coniospora secretome

To be able to act as virulence factors, many proteins, for example chitinases and proteases, need to be secreted. Some virulence factors are secreted into host cells, and can be targeted to specific organelles. We therefore complemented the InterproScan and PHI-base analysis with a focused and more thorough *in silico* investigation of the *D. coniospora* secretome. A total of 608 proteins (7%) were predicted with high confidence to be secreted (Fig 6 and S11 Table). They included the Saposin A domain protein g3895.t1, as well as 6/10 M35 domain proteins. More than a third of the putative secreted proteins (242/608), including 5 of the 6 secreted M35 domain proteins, and multiple proteins containing several different glycosyl hydrolase domains (GH2, 3, 16, 18, 20, 31, 35 47 and 65; see above), were also predicted to target a host cell organelle (e.g. nucleus or mitochondria), and of these 27 were homologous to proteins present in PHI-base with a demonstrated role in virulence (Fig 6 and Tables 7 and S11). In addition to chitinases, among the 27 predicted proteins, there were alkaline, aspartic, metallo-, subtilisin-like and cuticle-degrading proteases, all of which potentially contribute to the destruction of host tissue. There were also heat-labile enterotoxin homologues that would

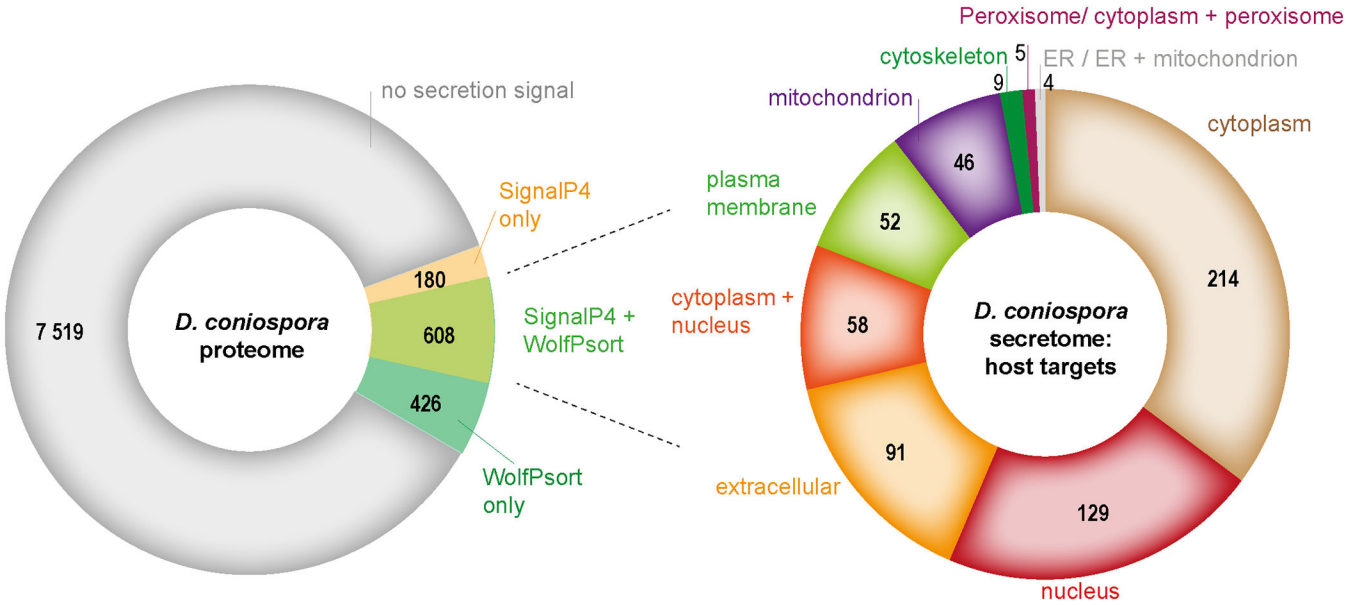


Fig 6. Predicted secreted proteins in *D. coniospora*. The left hand chart shows the distribution of protein predicted to be secreted by 2 different computational methods. For the proteins predicted to be secreted by both, the right hand chart indicates the predicted sub-cellular localisation.

doi:10.1371/journal.pgen.1006017.g006

Table 7. Potential virulence factors among *D. coniospora* proteins predicted to be secreted and targeted to a host organelle.

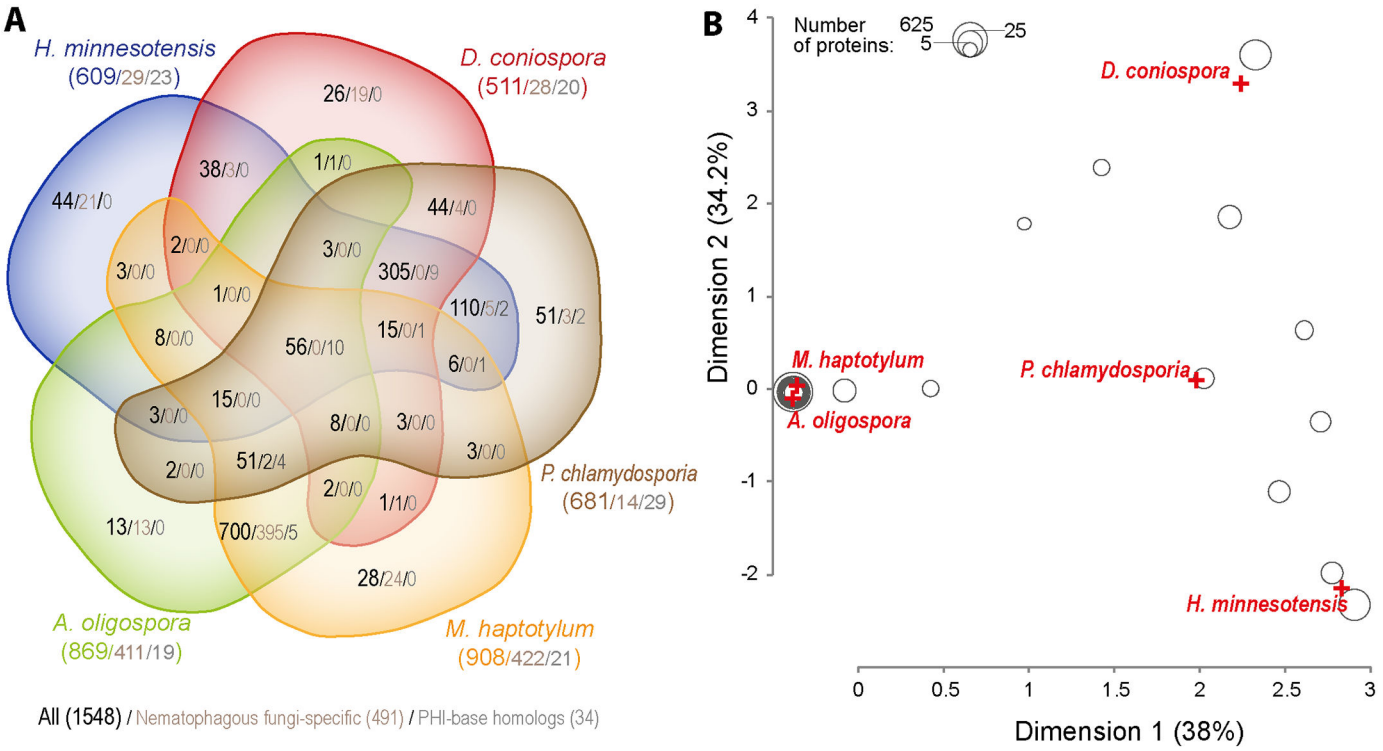
Gene ID(s)	PHIBASE ID	BLASTP hit	Brief description
g3389.t1	PHI:434	GEL1_ASPFC	1,3-beta-glucanosyltransferase
g6911.t1	PHI:2117	ORYZ_ASPCL	Alkaline protease
g1715.t1	PHI:73	YPS3_YEAST	Aspartic proteinase
g2067.t1	PHI:1046	SOL5_ALTSO	Bifunctional solanapyrone synthase
g2448.t1	PHI:383	SOD5_ARTBC	Cell surface Cu-only superoxide dismutase
g375.t1, g3187.t1	PHI:144	CHI1_APHAL	Chitinase 1
g2033.t1	PHI:2117	CUDP_METAN	Cuticle-degrading protease
g3550.t1	PHI:2570	CYB2_WICAO	Cytochrome b2, mitochondria
g858.t1	PHI:3078	E9DYG9_METAQ	Endonuclease/exonuclease/phosphatase family protein
g656.t1, g6818.t1	PHI:479	MEP1_COCP7	Extracellular metalloprotease
g960.t1	PHI:1071	GLU2A_SCHPO	Glucosidase 2 subunit alpha
g964.t1	PHI:698	E2AA_ECOLX	Heat-labile enterotoxin IIA
g496.t1	PHI:698	E2BA_ECOLX	Heat-labile enterotoxin IIB
g1983.t1	PHI:2920	ASO_CUCPM	L-ascorbate oxidase
g2355.t1, g5020.t1	PHI:785	MU157_SCHPO	Meiotically up-regulated gene 157 protein
g6255.t1	PHI:184	PRY2_YEAST	Pathogenesis-related protein 2
g4517.t1	PHI:1071	AGDC_ASPFU	Probable alpha/beta-glucosidase agdC
g1115.t1	PHI:68	OPSB_ASPOR	Probable aspartic-type endopeptidase
g2153.t1	PHI:184	PRY1_ARTBC	Probable pathogenesis-related protein
g5968.t1, g6831.t1	PHI:2654	A2965_ARTBC	Putative amidase
g1191.t1	PHI:1166	ATG15_CHAGB	Putative lipase (Autophagy-related protein 15)
g4757.t1	PHI:2117	SUB2_PSED2	Subtilisin-like protease 2
g7479.t1	PHI:891	G4MVB6_MAGO7	Uncharacterized protein

doi:10.1371/journal.pgen.1006017.t007

similarly be predicted to be direct effectors of virulence or play a role in *D. coniospora*'s interactions with other microbes (see below).

Comparative analysis of nematode-destroying fungi

The above results illustrate how secreted proteins can be key to virulence. To investigate commonalities and differences in the molecular basis of nematode infection, we therefore conducted a comparative analysis of predicted secretomes between the 12 fungal species, focusing on the 5 that are nematopathogenic. Using reciprocal BLASTP analyses, we first determined high-confidence clusters of orthologous proteins among the 12 species. We then concentrated on the 1548 clusters containing at least one protein from a nematopathogenic species (S11 Table), and calculated their distribution across the 5 species (Fig 7A). While there was a substantial overlap between *A. oligospora* and *M. haptotylum*, with 395/700 shared clusters unique to these 2 species, very few of the clusters uniquely shared between *D. coniospora* and *H. minnesotensis* or *P. chlamydosporia* were restricted to nematopathogenic species (3/38 and 4/44, respectively). Indeed, there were only 9 clusters present in *D. coniospora* and another nematopathogenic fungus but not any of the 7 non-nematopathogenic species. Only one of these clusters corresponded to proteins with a conserved domain, namely fungal hydrophobin (PF01185), also found in rodlet proteins, a major component the hydrophobic sheath, or rodlet



C

```
g2511.t1 MKLATTCSYAVTLLVGPATASLIGSESFFVKNNLLTAPDFGMLRLELASCLRGA
g2509.t1 MKLATTSSYAVTLLVGLVTASPL-ADSSLAKRGLIPPSEFQKEQSELASCLERF
g2507.t1 MKFTTTYSYAATLLVGLATASPL-ADSSVARRGGFTASDFLYERVLLVRCLEDF
g2506.t1 MKFATTYSYAVAILVGLATASPL-ADPSVARRSVLTASDFQDERVLLVGCLEDS
g2510.t1 MKF---CSYAVS-LVGLATASPL-ADSSLAGRDMTAASDFQDERVQLVDCLERF
g2508.t1 MKFATTCSYAVTLLVGLVTASPL-ADSSLAGRGMMLASDFKDERVQLVSCLEHF

g2511.t1 CQPASGITLFGTWAISQVNRIKSEADLNQALKECRGGCIRGIVSDVEM--GTKKKC
g2509.t1 CQPANGVR-HSGWATGQVSGINNKADLEQGLKECKRGCIRGIVKNVQKPEEGKSVDA
g2507.t1 CQ-PDGRAALSNWATSLVSWVD-KADLDQALKECKRGCTQGVLSGIKP--QS-----
g2506.t1 CQ-PDGPAAFSNWATSLVSRIEGKADLDKALKECKRGCVRGVLSGIGP--QS-----
g2510.t1 CQLAHGPAVHGGWATLQVSWIKSRAELDQALKECKRGCVRGVVRNVKK--KLSWKVK
g2508.t1 CQAADGLPAHSAWATSQVDRIKSKADLDQALKECKRGCVRGILSDVEK--SQ-----
```

Fig 7. Comparative analysis of the predicted secretome of *D. coniospora*. (A) Distribution of sequence-based protein clusters across five nematopathogenic fungal genomes. Only clusters containing at least one secreted protein are shown. Except for empty sectors, in each sector, there are 3 numbers: total number of clusters/number of clusters with members only in nematophagous fungi/number of clusters with at least one member that matches a PHI base entry with an annotation of reduced virulence or loss of pathogenicity. (B) A correspondence analysis of sequence-based clusters of secreted proteins from five nematopathogenic fungi. The first two dimensions are shown. Crosses represent the position of the individual fungal species and circles represent protein clusters. Circles are sized according the number of constituent proteins as indicated. When clusters have identical coordinates, the size of the circle represents the sum of the number of proteins in each cluster. For example, the circle at (2.33, 3.6) corresponds to 19 clusters of proteins, in this case unique to *D. coniospora*, including Cluster01087. The proximity of each circle to the species' apices is a measure of the contribution of the species to that cluster's content. The distance between the circles is a measure of the similarity of their content (number of proteins from each species). (C) Multiple sequence alignment of proteins from the *D. coniospora*-specific cluster Cluster01087. Only 3 of the 6 proteins are predicted to be secreted (g2506.t1, g2508.t1, g2511.t1; S11 Table).

doi:10.1371/journal.pgen.1006017.g007

layer, that covers the surface of fungal spores, required in *Aspergillus nidulans* for efficient spore dispersal [70].

To gain a more synthetic and rigorous overview of these results, we used correspondence analysis of the clusters of the secreted proteins from the 5 nematopathogenic fungi. Correspondence analysis is conceptually similar to principal component analysis, but can be used with

categorical rather than continuous data. The first 2 dimensions explained more than 70% of the observed distribution and separated the species. As expected from the numerical overlaps of protein clusters, *A. oligospora* and *M. haptotylum* were much closer together, and furthest from *D. coniospora* and *H. minnesotensis* (Fig 7B). The separation of the latter 2 species reflects their unique clusters. For *D. coniospora*, the most populated cluster of proteins (Cluster01087) contains 6 members (Fig 7C). The corresponding genes (g2506.t1-g2511.t1) are together in the genome. While the 6 proteins have no annotation or predicted function (S5 Table), 5 other clusters include at least one member with a PFAM domain. Consistent with our previous domain-centric analyses, among them we found the M35 (PF02102) and heat-labile enterotoxin alpha chain (PF01375) domains, reinforcing the notion that these domains potentially characterize the pathogenic capacity of *D. coniospora*. The remaining 3 were a domain of unknown function (PF11999), and 2 expected to be involved in fungal adhesion, the GLEYA (PF10528) domain, [71], and the hydrophobic surface binding protein A (PF12296) domain discussed in the next section. It is noteworthy that GLEYA domain proteins are highly expressed by nematode-trapping fungi during infection of *C. briggsae* [5].

Stage-specific gene expression

While analysis of the genome sequence reveals predicted proteins putatively involved in virulence or other aspects of fungal physiology, it is necessary to establish when the corresponding genes are in fact expressed. To gain a first insight into the genes potentially important at different stages of the life cycle of *D. coniospora*, we chose to compare gene expression between mycelia that had been grown in liquid for several generations in the absence of nematodes, and spores harvested from infected *C. elegans* and starting to germinate *in vitro* (S2 Fig). These two morphological forms were expected to provide a broad though not necessarily exhaustive repertoire of expressed transcripts. The combined set of 48.6 million paired-end reads were mapped to the set of predicted genes. Inspection of the most highly expressed genes (arbitrarily the top 50) in mycelia and spores, revealed a substantial number of genes encoding basic metabolic enzymes, such as glyceraldehyde 3-phosphate dehydrogenase, as well as multiple proteins involved in translation (ribosomal proteins and elongation factors) and protein folding (chaperones of the heat shock protein family), presumably reflecting the need for protein synthesis during growth. We then used stringent criteria to define a small set of genes that were differentially regulated between germinating spores and mycelia, (61 and 86 genes, respectively; S12 Table). The list of spore-specific genes was particularly interesting. It included 16 encoding predicted proteins with no identified PFAM domain or homolog in UniProtKB. Several are found in the secretome-associated and/or orthoMCL-defined clusters described above including g2508.t1 and g2509.t1 (group 6 and Cluster01087), g4179.t1 (group 2), and g2083.t1 (group 34). The functional role in spores of the members these different paralogous groups merits further investigation in the future.

Two genes g3607.t1 and g6474.t1 encode proteins that contain the hydrophobic surface binding protein A domain (PF12296) mentioned above. By analogy with the eponymous protein from *Aspergillus oryzae* [72], we hypothesize that this protein is important for spore attachment to the cuticle of *C. elegans* and that further, they may act together with g5675.t1 that contains a CFEM domain (PF05730 [69]), that is structurally related to fungal adhesins and is highly preferentially expressed in spores. The differential expression of such genes will depend on stage-specific expression of transcription factors, such as g5153.t1 that is predicted to be a transcription factor with Zn_2Cys_6 (PF00172) and fungal-specific (PF11951) domains, and is also preferentially expressed in spores (S12 Table).

Genes that are more highly expressed in mycelia than spores would be predicted to be important for vegetative growth but also potentially for virulence. The list of these genes included candidates in both categories (S12 Table). Thus, on the basis of the domains found in the corresponding predicted proteins, 5 genes are associated with carbohydrate metabolism (GO:0005975; g1835.t1, g1896.t1, g3555.t1, g7200.t1, g7950.t1), while g7950.t1 corresponds to a pyruvate/2-oxoglutarate dehydrogenase, a key metabolic enzyme, and fg7260.t1 to a sulfide: quinone oxidoreductase that catalyzes the first step in the mitochondrial metabolism of H_2S. These are expected to fulfil metabolic needs that are not present during the early growth of spores.

Otherwise, there is g3889.t1 that encodes a highly conserved protein, annotated as a putative NRPS-like enzyme in multiple fungal species; it was not identified as such by SMURF analysis. g3999.t1 encodes a subtilisin and was also preferentially expressed in mycelia. Subtilisin-type proteases are associated with virulence, including nematicidal activity [73], in many species [6]. As in other nematophagous fungi [8], as detailed above, *D. coniospora* has a large family of subtilisin-type proteases (PF00082; S7 Table and Table 5). These different genes are all potentially linked to growth in the nematode host. Strikingly, 46 of the 86 (53.5%) proteins corresponding to genes preferentially expressed in mycelia currently have no identifiable domains nor homologs in the UniProtKB database, compared to 30.8% for the full set of 8733 predicted proteins (S5 and S12 Tables), and among them 41/46 are shorter than the median length. Manual searches suggest that many may have homologues in other fungal species. As a single example, fg6211.t1 is expressed at almost 10-fold higher levels in mycelia than spores, encodes a predicted 70 amino acid protein and matches a predicted 64 amino acid protein of unknown function from *Trichoderma reesei* (Genbank XP_006962079). Determining the role of these different genes will require extensive functional analyses in the future.

Gene expression during infection of *C. elegans*

While this analysis revealed genes potentially involved in virulence, an important question is what fungal genes are actually expressed during infection. Having the annotated *D. coniospora* genome in hand allowed a re-examination of RNAseq data obtained from samples of *C. elegans* infected by *D. coniospora* (NCBI SRA SRX036882 and [74]). From sequencing of samples taken 5 and 12 h post-infection (p.i.), a small number of reads among those that did not align to the *C. elegans* genome could be aligned to predicted *D. coniospora* genes. Together, 537 gene models were covered by at least one read, with 339 only at 5 h (p.i), 142 only at 12 h (p.i.) and 56 at both time-points (S13 Table).

Focusing on genes for which there were at least 3 matching reads (S13 Table), as might be expected, many corresponded to genes that were highly expressed in mycelia and/or spores (30/47 within the top 15 percentile for expression; i.e. >1192 and >1298 reads, for mycelia and spores respectively; S12 Table). In addition to 14 ribosomal genes, they included the CFEM domain protein, g5675.t1, mentioned above. Six encode proteins homologous to ones present in PHI-base with a demonstrated role in virulence. For example, g6659.t1 corresponds to a component of the mitochondrial membrane ATP synthase complex (S13 Table). As with the other such genes, its role in virulence probably reflects a general function in fungal physiology and growth.

For the remaining 17 genes that were not highly expressed in mycelia and/or spores, BLASTP searches at NCBI lead to the identification of potential homologs for 15 of the corresponding predicted proteins (e-value $<10^{-10}$), across a range of species (Table 8). Among them, 3 encoded homologues of PHI-base listed virulence factors (S5 Table and Table 7). The first, g2153.t1 (PHI:184), potentially encodes a cysteine-rich secretory protein family

Table 8. Selection of *D. coniospora* genes expressed during the infection of *C. elegans*.

	Read counts					Best Genbank BLASTP hit			PFAM domains	
Gene ID	5 hour p.i.	12 hour p.i.	Mycelia	Spores	Length (a.a.)	.Genbank	Length (a.a)	Species	ID	Name
g2153.t1	8	0	81	189	307	GI:302895657	324	*Nectria haematococca*	PF00188	Cysteine-rich secretory protein family
g2389.t1	3	0	40	29	118	GI:799240667	128	*Hirsutella minnesotensis*	-	
g2390.t1	4	0	57	34	303	GI:667661323	308	*Beauveria bassiana*	PF11327	Protein of unknown function (DUF3129)
g261.t1	0	4	411	167	69	GI:952549466	69	*Aspergillus lentulus*	PF11034	Protein of unknown function (DUF2823)
g3607.t1	7	2	147	593	428	GI:908391879	373	*Tolypocladium ophioglossoides*	PF12296	Hydrophobic surface binding protein A
g4094.t1	10	11	26	73	111	-	-	-	-	
g4805.t1	3	0	123	82	323	GI:666406534	327	*Stachybotrys chartarum*	-	
g4815.t1	3	1	3	35	354	GI:908393692	353	*Tolypocladium ophioglossoides*	PF11790	Glycosyl hydrolase catalytic core
g4816.t1	3	2	0	54	343	GI:949387848	548	*Rosellinia necatrix*	PF12430	Abscisic acid G-protein coupled receptor
g4817.t1	4	0	3	262	446	-	-	-	-	
g5054.t1	3	0	23	111	363	GI:969886810	403	*Trichoderma gamsii*	-	
g5671.t1	1	5	422	198	163	GI:799246947	108	*Hirsutella minnesotensis*	-	
g6474.t1	10	0	40	426	331	GI:389629282	257	*Magnaporthe oryzae*	PF12296	Hydrophobic surface binding protein A
g6475.t1	3	0	89	166	600	GI:743663295	600	*Metarhizium guizhouense*	PF01532	Glycosyl hydrolase family 47
g6844.t1	2	1	77	159	160	GI:531863496	231	*Ophiocordyceps sinensis*	-	
g6908.t1	2	1	5	4	563	GI:667643613	395	*Beauveria bassiana*	PF00082/ PF05922	Subtilase family/Peptidase inhibitor I9
g7182.t1	3	0	107	161	416	GI:672821344	241	*Mortierella verticillata*		

The Gene IDs in bold indicate genes that are neighbours in the genome. Those that are underlined are homologous to proteins in PHI-base annotated as being important for virulence.

doi:10.1371/journal.pgen.1006017.t008

(PF00188) member, similar to the pathogen-related protein Pry1. Members of this family have different roles in various pathogenic fungal species, including the neutralization of host defenses, and antimicrobial activity to inhibit the growth of competing microorganisms [75]. The second, g2390.t1 (PHI:257), encodes a widely conserved cell surface protein of unknown function, and the third, g6908.t1 (PHI:2117), an alkaline serine protease, homologous to the peptidase S8 of *Beauveria bassiana*. Interestingly, among the 17 genes, there were 3 groups of neighbours, suggesting a coordination of gene expression at the genome level. These included a group of 3 genes, g4815.t1–g4817.t1, that encode proteins that are completely unrelated in sequence: g4815.t1, one of 3 *D. coniospora* GH128 proteins, from a recently described glycoside hydrolase family ([76]; S8 Table), g4816.t1, a G-protein coupled receptor, and g4817.t1 that does not currently have homologues in any other species. For the other predicted proteins, it is notable that 2 encode proteins with a hydrophobic surface binding protein A (PF12296) domain, also found in chitinases (e.g. KID89971). One of these, g6474.t1, was mentioned above since it is preferentially expressed in spores, and may be important for the initial adhesion to and penetration of the nematode cuticle. This is consistent with the fact that the corresponding RNAseq reads were only found at the early time-point of infection.

Transformation and genome modification

The different *in silico* analyses reported above led to the identification of a very large number of candidate virulence genes. Addressing their functional importance would be greatly facilitated by the availability of techniques for the genetic transformation of *D. coniospora* and for targeted editing of its genome. By screening a number of different liquid media (C. Couillault, personal communication), we found that *D. coniospora* grew well in a rich, cholesterol-supplemented medium. We used fresh liquid cultures of *D. coniospora* mycelia to generate protoplasts, which were then transformed using a standard technique of polyethylene glycol (PEG)/$CaCl_2$-mediated DNA uptake [77]. As a proof of principle, we transformed protoplasts with a plasmid (pLH4237) in which expression of a gene encoding a chimeric hygromycin B phosphotransferase::GFP protein [78] was driven by the *D. coniospora* ß-tubulin promoter (*ß-tub*p::HPH::GFP). The resultant recombinant fungus exhibited hygromycin resistance and strong GFP fluorescence in both spores and mycelia (Fig 8A).

To test the possibility of specifically knocking out a gene's function, we chose to target the *D. coniospora* homolog of the *so* (*soft*) gene (NCBI Gene ID: 3880225), required in other fungi for anastomosis (mycelial fusion) since this was expected to give a clear viable and visible phenotype [79], while at the same time not greatly altering virulence [80]. We therefore flanked our *ß-tub*p::HPH::GFP construct with arms homologous to the 5' and 3' regions of the *D. coniospora so* gene (*Dso*; g1469.t1) and used this construct (in pLH4256) to transform protoplasts. We obtained hygromycin-resistant GFP-expressing transformants in which, as demonstrated by PCR (S3 Fig) and sequencing, the *Dso* gene was replaced by the *ß-tub*p:: HPH::GFP cassette. The mutant exhibited the expected anastomosis defect. In contrast to the wild-type strain (Fig 8A and 8B, S2B and S2C Fig), neighbouring mycelia were never observed to fuse *in vitro* (Fig 8C) or during infection of *C. elegans* (Fig 8D and 8E). We have therefore the capacity to make targeted modifications of the *D. coniospora* genome.

Having fungal strains that express a fluorescent protein opens many new possibilities for future research. One immediate consequence is that we were able to follow the infection *in vivo* directly using fluorescence microscopy (Fig 8A, 8B, 8D and 8E). We also wondered whether this would offer a new way to quantify the progression of the infection. Fluorescence in *C. elegans* can be measured *in vivo* using the Complex Object Parametric Analyzer and Sorter (COPAS) Biosort. When the Biosort's Profiler is used, as well as a single measurement for each

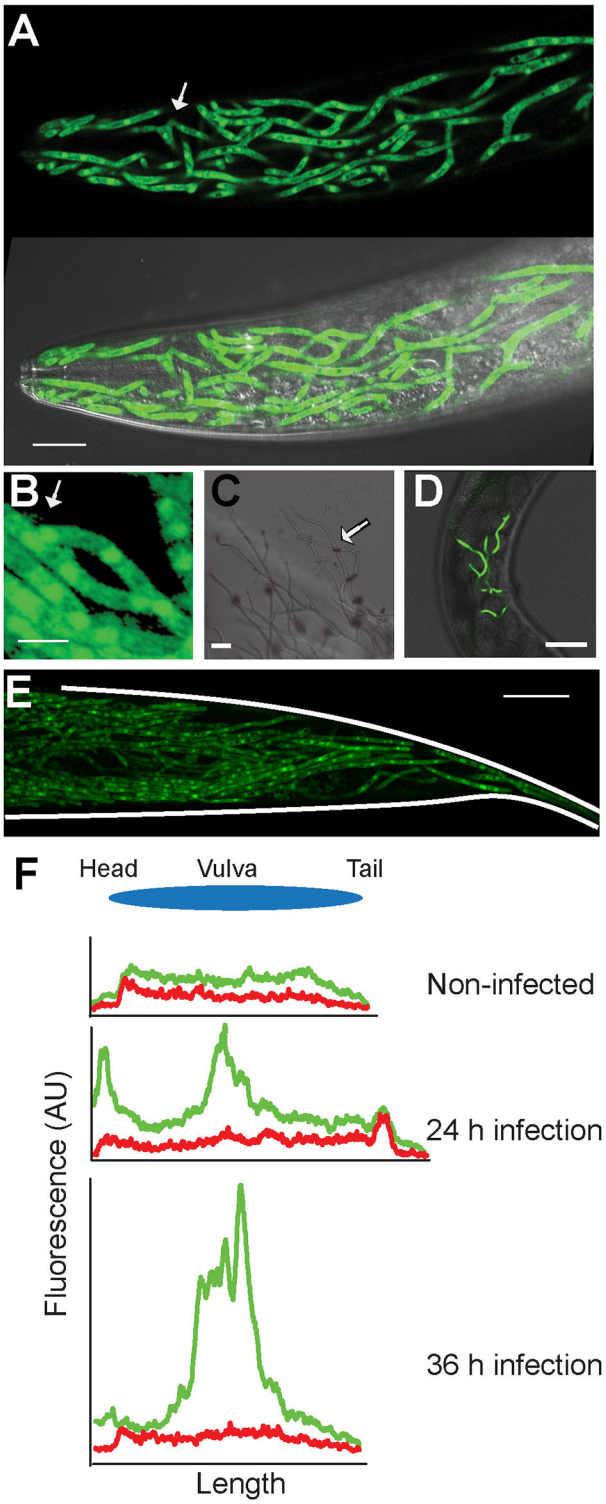


Fig 8. Transformation of *D. coniospora*. (A) A recombinant strain of *D. coniospora*, expressing GFP under the control of a ß-tubulin promoter, viewed by fluorescence microscopy (upper panel) combined with differential interference contrast microscopy (lower panel; scale bar, 20 μm). The fusion of 2 mycelia is highlighted by an arrow. (B) Higher magnification view of fused mycelia (scale bar, 5 μm). (C-E) The *Dso* mutant has a defect in anastomosis. (C) As highlighted with the arrow, in culture, mycelia are seen to grow across one another but never fuse. The mutant strain was engineered to express GFP constitutively. Fungal

mycelia growing in living animals, visualized using a stereo fluorescence dissecting microscope (D), or confocal fluorescence microscope (E; the shape of the worm is traced by white lines), were not observed to fuse. Worms had been infected overnight (D) or for 60 h (A, B, E) before images were taken. Scale bars in C, D, E: 20, 50, 40 μm. (F) The growth of the fungus can also be followed using the Profiler of the COPAS Biosort. The graphs show fluorescence profiles for green and red channels for an uninfected worm (top); a worm infected at the head and vulva (peak in green signal on the left and in the middle, respectively in the middle graph) and analysed after 24 h; another worm infected at the vulva (peak in green signal in the middle, bottom graph) and analysed after 36 h. Fluorescence and length are measured in arbitrary but constant units.

doi:10.1371/journal.pgen.1006017.g008

worm, one obtains a readout of fluorescence intensity along the length of the worm [81]. We found that the COPAS Biosort was sufficiently sensitive to allow us to follow the progression of the infection in a qualitative (Fig 8F) and potentially quantitative manner.

Initial characterisation of a potential fungal counter-defensive strategy

Our *in silico* analysis highlighted the atypical presence of 2 proteins with saposin A (PF02199) domains, one of which (g3895.t1) was predicted to be secreted (S11 Table). Mammalian saposins are synthesized as precursor proteins (prosaposin) that contain four Saposin-B domains (PF05184; PF03489) and two Saposin-A domains that are removed during the process of activation. Saposin-B domains also occur in proteins without Saposin-A domains across many species including nematodes. In *C. elegans*, they are represented by a large family of 23 proteins, also called caenopores [82, 83]. They are structurally similar to the innate defense proteins of the SAPLIP family, including vertebrate NK-lysin and granulysin [84]. Several of them are upregulated upon infection, some by multiple pathogens ([85] reviewed in [22]), including *spp-2*, *spp-6*, *spp-13*, *spp-14* and *spp-15* that are induced upon infection by *D. coniospora* [74]. A number of the SPP caenopores/saposins have been demonstrated to play a role in host innate immunity [86, 87] suggesting that certain SPP proteins could be also be direct effectors of antifungal defense.

The *D. coniospora* protein g3895.t1 is characterized by the presence of 3 Saposin-A domains, but no Saposin-B domain. There are currently no clear orthologs in any species in publically available databases. Given its unusual structure, we hypothesised that this protein, which we call here SapA, might act as an inhibitor of one or more nematode Saposin-B domain-containing caenopores/saposins and thereby interfere with host defense.

As a first step in the analysis of *sapA*, we chose to assay directly its expression during the infection of *C. elegans*. Using the *D. coniospora* actin gene as a control, by RT-PCR we observed a clear increase in the relative level of expression of *sapA* across the time-course of infection (S4A Fig). To define *in vivo* the spatio-temporal expression pattern of the corresponding protein, we made use of our capacity for transformation to produce recombinant fungus expressing the SapA protein tagged with dsRed at its C-terminus (SapA::dsRed), under the control of its own promoter. Strong dsRed expression was observed at the surface of spores, but not on mycelia early in the infection. Expression was then seen at the tips of growing hyphae at the moment when they approached the apical surface of the epidermis, before penetrating the cuticle from the inside (Fig 9A).

To test the hypothesis that the *D. coniospora* SapA might interact physically with one or more of the *C. elegans* caenopores/saposins, we incubated an extract of proteins purified from the *D. coniospora* strain expressing SapA::dsRed with 3 different purified recombinant *C. elegans* SPP proteins [83, 87], each possessing a C-terminal His-tag. The SapA::dsRed was then immunoprecipitated, together with any bound SPP protein, using an anti-dsRed antibody. To probe for a possible interaction with the SPP proteins, the immunoprecipitated material was analysed by Western blotting, using an anti-His-tag antibody. Although the 3 samples had

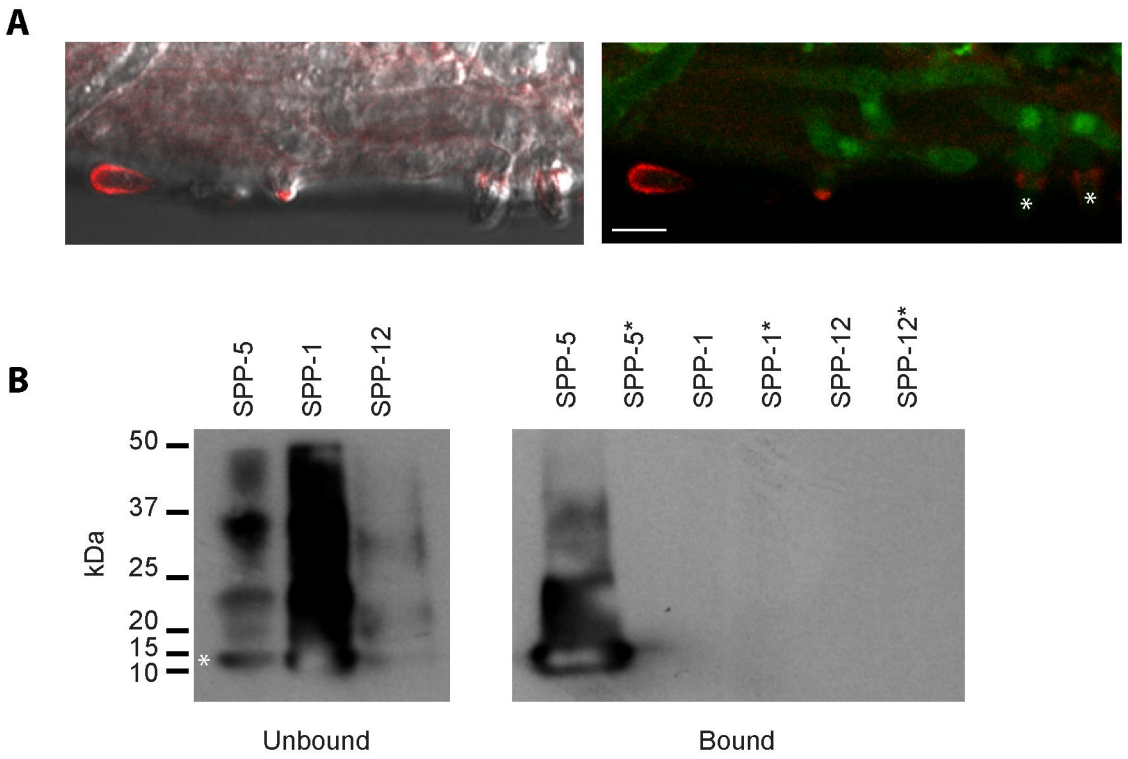


Fig 9. Saposin A-domain protein expression during infection of *C. elegans* and its *in vitro* interaction with SPP-5. (A) A strain of *D. coniospora* engineered to express GFP constitutively and a SapA::dsRed chimeric protein under the control of the *sapA* promoter, visualized using confocal fluorescence microscope (right panel) combined with differential interference contrast microscopy (left panel), during the infection of *C. elegans*. In both panels, a bright red fluorescent spore can be seen on the left. In the centre, a mycelium that is starting to exit the worm shows bright red fluorescence at its tip. On the right, 2 adjacent mycelia that have emerged can be seen. At the point where they leave the epidermis, a ring of SapA::dsRed can be seen, marked by asterisks in the right-hand panel. A general, less concentrated, signal can also be seen in the infected tissue. Worms had been infected for 60 h before images were taken. Scale bar = 5 μm. (B) Physical interaction between SPP-5 and SapA::dsRed. Recombinant His-tagged SPP-1, SPP-5 or SPP-12 was mixed with a protein extract from fungi expressing SapA::dsRed. The mix was analysed by Western blot probed with an anti-His-tag antibody before (left hand panel) or after (right hand panel) immunoprecipitation with anti-dsRed antibody-coated beads. In the sample before immunoprecipitation, in addition to the band at the expected size (11.9 kDa) marked by an asterisk, higher molecule weight species were detected, corresponding to the previously described oligomerization [86]. SPP-5 was co-immunoprecipitated with SapA::dsRed, principally in its monomeric form, but not if incubated with blocked beads (lanes marked with an asterisk; a control for non-specific binding). Neither SPP-1 nor SPP-12 gave any indication of co-immunoprecipitating with SapA::dsRed, even if SPP-1 was more abundant in the sample before immunoprecipitation.

doi:10.1371/journal.pgen.1006017.g009

been incubated with equal quantities of fungal protein, while there was no indication of any interaction between SapA::dsRed and SPP-1 or SPP-12, we observed a clear co-immunoprecipitation of SapA::dsRed and SPP-5 (Fig 9B). Interestingly, SPP-5 is markedly more closely related to SPP-2 (79% similar) and the other infection-induced SPPs than are either SPP-1 or SPP-12 (S4B Fig).

Discussion

D. coniospora was first described 75 years ago as a parasite of *Rhabditis* nematodes in leaf mold, and then named *Meria coniospora* [88]. Its spores adhere to many different nematode species, and it is capable of infecting a relatively broad range of hosts including *C. elegans*, soybean cyst and root knot nematodes [14, 89–93]. While the different steps in the infectious process have been documented at the ultrastructural level [10, 94], nothing is known at the

molecular level. Here, we have provided a first annotated genome sequence of *D. coniospora* that will serve as a starting point for future functional studies, as well as for more refined predictions of gene structure (S1 Text).

To make a comparative study of the nematophagous lifestyle, we chose 4 fungal species that use different strategies to infect nematodes. Like *D. coniospora*, *H. minnesotensis* is an endoparasitic fungus of the family Ophiocordycipitacea. It naturally infects soybean cyst nematodes using non-motile spores. Although it has never been reported to be one of its natural pathogens, it can infect *C. elegans* in the laboratory [9, 95]. *M. haptotylum* and *A. oligospora* are phylogenetically distant species that infect diverse nematodes, including *C. elegans*, after trapping them with their adhesive knobs and nets, respectively. *P. chlamydosporia* lies between these 2 pairs phylogenetically, infects nematode eggs, and in common with the nematode trapping fungi can grow as a saprophyte [3, 4, 6, 96–102].

Our phylogenetic analysis confirms the assignment of *D. coniospora* to the family Ophiocordycipitacea. It is interesting to note that other members of the Ophiocordycipitaceae family, *Haptocillium sphaerosporum* and *Harposporium spp*. are also pathogens of *C. elegans* [20, 74]. Among sequenced fungi, *D. coniospora*'s closest relative is currently the truffle-parasite *Tolypocladium ophioglossoides*. Our analysis thus provides a further illustration of the polyphyletic nature of nematode parasitism and will contribute to the on-going debate regarding the acquisition of host specificity in pathogenic fungi [6, 67].

One factor that can contribute to the emergence of virulence traits is the amplification of transposable elements that facilitate genome rearrangements. The *H. minnesotensis* genome represents an extreme example since 35% of its genome was reported to be composed of transposable elements [9]; in *D. coniospora* they make up less than 2%. *D. coniospora* also has a substantially reduced range of glycoside hydrolase (GH) enzymes compared to the nematode-trapping fungi, in direct contrast to the extensive repertoire of *P. chlamydiosporia* [98]. This presumably reflects a decreased capacity to adapt to diverse environments [103]. Unlike *P. chlamydiosporia* that can infect plants as well as nematodes, *D. coniospora* is an obligate nematode-specific parasite [104]. This slimming down of the genome is reflected in diverse other protein families. Thus, in addition to the different families described above, for example, there are few multi copper oxidases and lytic polysaccharide mono-oxygenases, so-called "auxiliary activities" linked to lignocellulose conversion [105]. Its genome also contains fewer genes related to sugar/inositol transport, which are involved in the establishment of plant-fungus relationships in *M. anisopliae* [106]. Thus *D. coniospora* does not show the pattern of gene family expansions observed in characterized nematode-trapping fungi, which are more similar to those seen in plant pathogens than to insect and animal pathogens [5].

An atypical loss from the *D. coniospora* genome is of genes encoding NACHT-domain proteins, which as mentioned above are present in multiple copies in all the 11 other fungal species analysed. The NACHT domain is a constituent domain in one of the two main classes of NOD domain proteins. The second class has the NB-ARC (PF00931) domain, a signalling motif shared by plant resistance gene products and regulators of cell death in animals [107]. These too are absent from the predicted set of *D. coniospora* proteins. Again this is highly unusual, since they are found often in all the other species, (8, 9, 10 and 15 times in *M. haptotylum*, *A. oligospora*, *P. chlamydosporia* and *H. minnesotensis*, respectively). Indeed, NOD domain proteins are present broadly among fungi and have important roles in fungal non-self recognition and in defence systems [60]. Determining the reasons and consequences of this intriguing loss of NOD domain proteins from *D. coniospora* remains a challenge for the future.

On the other hand, both insect and nematode parasitism requires a broad set of genes involved in detoxification and resistance to oxidative stress, such as glutathione S-transferases, cytochrome P450 genes, and carboxylesterases [40]. These are present in the *D. coniospora*

genome in similar numbers to its entomopathogenic neighbors. More than 100 protein kinases were detected. Many are likely to regulate the infection process, as previously described, for example, for a *M. anisopliae* protein kinase A [108] for which there is a direct *D. coniospora* ortholog (g2806.t1). They may in turn be regulated by different families of transcription factors, including the Zn_2Cys_6 fungal-type, bZIP and bromodomain-containing families, which are all well-represented in the *D. coniospora* genome, and are often involved in the control of gene expression linked to virulence [109].

As previously observed for other nematophagous species, the specialization of *D. coniospora* is also reflected in the high degree of structural innovation in its predicted proteome. Thus, for example, in *H. minnesotensis*, 11% of proteins were described as species-specific and lacking any recognisable domain [9]. In *D. coniospora* the corresponding figure is currently 16.3%. With increased sampling of closely related species, these figures will drop in the future. Nevertheless, in addition to the many readily predictable virulence factors, there is a wealth of novel biology to be explored using *D. coniospora*.

Regarding the broad spectrum of potential virulence effectors, certain merit further discussion. The *D. coniospora* genome is predicted to encode multiple heat-labile enterotoxins. The presence of bacterially-derived enterotoxins in entomopathogenic fungi such as *B. bassiana* already presented a conundrum as these species are assumed to lack *per os* infectivity [53]; the same is true for *D. coniospora*. Among them, at least 10 are predicted to be secreted (Tables 4 and S6 and S9), so may conceivably be delivered into the host cytoplasm where they would be expected to perturb cellular homeostasis. An alternative explanation is that they may play a role in *D. coniospora*'s interactions with other microbes. Competition has already been reported between *D. coniospora* and *A. oligospora* during nematode infection [110]; antagonistic effects are likely to exist with other species and could rely on enterotoxin production.

Subtilisin-like serine proteases were mentioned several times above. This family has been linked to fungal virulence in nematophagous fungi in some studies (e.g. [8, 73, 111]). Consistent with such a link, subtilisin genes are highly expressed by *A. oligospora* and *M. haptotylum* during infection of the nematode *C. briggsae* [5]. There are 29 genes predicted to encode subtilisins in *D. coniospora*. This family is, however, well represented in all the fungal species analysed. For the 11 other species, the number ranges from 12 to 56, with a median value of 31. Since related pathogenic and non-pathogenic species can show the same type of gene expansion, and thus gene number is not correlated with pathogenicity, it has been suggested that the number of serine proteases in a species is related primarily to their role in digestion, whether or not the food source is dead or alive [112].

The genome also contains multiple cysteine-rich secretory protein family genes. These are frequently associated with fungal host adaptation or specialization [67]. As mentioned above, one was found to be preferentially expressed during infection of *C. elegans*, despite the poor coverage of the fungal transcriptome. This coverage was in fact remarkably low, despite deep sequencing, with only 0.0011% and 0.0012% of 77 million and 123 million RNAseq reads corresponding to *D. coniospora* transcripts from samples of worms infected for 5 and 12 hours, respectively. Clearly, reliable profiling of the *D. coniospora* transcriptome at early time points will require the development of methods to enrich fungal mRNA from samples of infected worms. Fortunately, several methods already exist (e.g. [113]); adapting them to this model will require further work.

Such an analysis is likely to be necessary to prioritize the overwhelming number of candidate virulence factors for in-depth functional study. This is particularly true in cases where gene families have expanded and where there is therefore the possibility of functional redundancy. Characterizing expression profiles can allow genes with non-overlapping patterns to be identified; they are less likely to be redundant. While we have shown that *D. coniospora* can be

readily genetically transformed, this remains a relatively time-consuming process that cannot be implemented on a large scale. Indeed, further work is also needed to develop functional assays to assess the role of the many genes potentially involved in pathogenesis or currently lacking any predictable role. The COPAS Biosort allows the analysis of hundreds of worms each minute. Having the possibility to transform *D. coniospora* with a fluorescence protein encoding reporter gene thus provides a non-invasive, non-destructive means to following infection at the level of individual worms at a large scale. It represents a first method of the type that will need to be applied to probe *D. coniospora* gene function during its infection of *C. elegans*.

We have also given one example of how tagging a protein fluorescently can constitute a first step in its functional characterization. We were able to determine the *in vivo* expression pattern for SapA::dsRed and demonstrate an *in vitro* interaction between it and the host antimicrobial effector SPP-5. We do not yet know whether this interaction mirrors an *in vivo* interaction. We currently favour the idea that SapA::dsRed will be capable of interacting with the host SPP caenopores/saposins that are up-regulated upon infection and thereby inhibit their antimicrobial activity. SPP-5 is very similar in sequence to SPP-2, SPP-6 and SPP-14 that are differentially regulated in *D. coniospora* infected worms. Clearly much work remains to be done to clarify the significance of these observations, including establishing which if any of the host SPP caenopores/saposins bind SapA *in vivo*, and determining whether this has any consequence for the progression of the fungal infection. Given the multiplicity of SPP proteins in *C. elegans*, this remains a substantial challenge for the future. Clearly, our analysis of the *D. coniospora* genome, and the tools we have developed, have opened many avenues for future investigation of this fungus's antagonistic interaction with *C. elegans*.

Materials and Methods

D. coniospora culture and nucleic acid purification

The *D. coniospora* strain ATCC 96282 (the kind gift of Hans-Börge Jansson; S1 Text) was either cultured at 25°C on solid Nematode Growth Medium (NGM) in the presence of *C. elegans* as previously described [114], or in liquid NGM with Yeast extract (NGMY see S1 Methods) in the absence of nematodes. For the extraction of fungal genomic DNA, roughly 10^9 fresh spores were inoculated in 100 mL NGMY medium. After 7 days culture, mycelia were harvested by vacuum filtering through a sterile 10 μm nylon membrane. The filter was placed in a 30°C incubator for around 3 h, until the mycelia were dry. The mycelia were then manually ground in a liquid nitrogen-cooled mortar. A 20 mg aliquot was transferred to a 1.5 mL tube and the DNA extracted using the DNeasy Plant Mini Kit (Qiagen), following the manufacturer's instructions.

To prepare samples for RNA extraction, fresh spores were first collected from 10 cm NGM plates of infected worms as previously described [114] and inoculated in 100 mL of NGMY liquid medium, and either cultured for 4 days before harvesting, or serially cultured (5 x 8 d cultures). For this, a 1 mL aliquot from an 8 d culture was strongly agitated for 5 min to disrupt the tight balls of mycelia and inoculated into fresh 100 mL NGMY liquid medium then cultured for a further 8 days. In both cases, samples were collected by filtering cultures through 0.22 μm Steritop units (Milipore), flash frozen in liquid nitrogen and stored at -80°C. Aliquots of 200 mg of spores or mycelia were dissolved in 1 mL Trizol (Invitrogen) in lysing matrix tubes (MP Biomedicals), homogenized for 20 s at 6 m/s in a FastPrep-24 (MP Biomedicals), incubated on ice for 2 min and homogenized a second time. RNA was purified using a standard protocol [115] and cleaned with an RNeasy mini Kit (Qiagen).

Library construction and sequencing

MIseq paired-end libraries were prepared from genomic DNA and sequenced on an Illumina MIseq sequencer as 2 x 150 bp paired-end reads following the manufacturer's standard procedures. For SOLiD sequencing, two mate-paired libraries were prepared from *D. conisopora* genomic DNA following the manufacturer's instructions (Mate-Paired Library Preparation user guide 4460958 Rev.B; Life Technologies, Carlsbad, USA). DNA was sheared to 1.5 kb or 3 kb fragments with Covaris System and Covaris Blue miniTUBES. Following nick-translation for 9.5 min, to generate fragments of $<$ 400 bp, the library was amplified through 14 rounds of PCR. Fragments between 250 and 300 bp were size-selected using 4% acrylamide gel and purified with SOLiD Library Micro Column Purification Kit prior to conversion for analysis with the 5500 WildFire system (Life Technologies, Carlsbad, USA).

cDNA library construction and Illumina sequencing of mRNA from spores and mycelia using an Illumina HiSeq 2000 platform was performed at the Beijing Genomics Institute (Shenzhen, China; http://www.genomics.cn/index.php) using their standard pipeline. More than 48 million 90 bp paired-end reads were obtained from each 200 bp insert library.

De novo genome and transcriptome assembly

MIseq reads were processed with BBDuk software, part of BBMap suite (http://sourceforge.net/projects/bbmap/) to filter out contaminants and low quality reads. The remaining reads were used for *de novo* genome assembly with Velvet [34], SPAdes [35], SOAPdenovo2 [36] and ABySS [37] using standard input parameters except for ABySS for which k values of 64 and 96 were used. The resulting assemblies were then scaffolded using SOLiD mate-paired reads. Only very high quality reads were used for this scaffolding step using SSPACE v2.0 [38] with the k parameter set to 5 (default value).

For use subsequent in gene prediction (see below) a *de novo* transcriptome assembly was performed on the combined sets of reads from the two sequenced libraries (from spores and mycelia) with Trinity [116] using default parameters.

Optical mapping

A whole genome map of *D. coniospora* was generated using the Argus Whole-Genome Mapping System (www.opgen.com). To obtain size-optimized restriction fragments (6–12 kb on average and no fragment larger than 80 kb across the genome) we used Enzyme Chooser (OpGen Inc., Gaithersburg, MD) that led to the selection of *Xba* I. We sorted out 42,622 molecules longer than 200 kb (average 310 kb) used for the assembly, performed with MapSolver software (www.opgen.com). The resulting map contigs were manually validated leading to 9 maps (ranging from 0.58 to 11.3 Mb) with a cumulative size of 31.8 Mb. These were used for comparisons with the *in silico Xba* I digestion profile of the scaffolds obtained after sequencing.

Functional annotation

Genome annotation was performed using standard open source software. Repetitive elements were mined using RepeatScout version 1.0.5 [42]. A specific repeat library was generated in order to mask the genome with RepeatMasker version 4-0-5 (Smit, Hubley, & Green, RepeatMasker Open-4.0; www.repeatmasker.org). TransposonPSI v08222010 (transposonpsi.sourceforge.net) was used to characterize different types of transposable elements. Non-coding RNA were identified using Rfam scan perl script v1.0 [117] and tRNAscan-SE v1.3.1 [43] for transfer RNAs.

Gene prediction and annotation

We performed a first round of gene prediction using Augustus [45], trained using the Trinity-derived *D. coniospora* transcripts to predict 8111 protein-coding genes. In a second round, we used Augustus trained using *Fusarium graminearum* (available from http://bioinf.uni-greifswald.de/augustus/) and retained 631 additional predicted proteins that were either (i) conserved, (ii) supported by RNAseq data or (iii) contained a PFAM domain. Inspection of the combined set led to the removal of 9 aberrant proteins (S1 Text), giving a final set of 8733 predicted proteins. Subsequent inspection of the alignment of the unassembled RNAseq reads and the Trinity-derived transcripts to the predicted gene models revealed occasional inconsistencies. Thus as with any first round genome-wide gene prediction and sequence annotation, some errors will need to be resolved in future versions of the genome (see S1 Text).

The set of Augustus-predicted protein-coding genes were annotated using Interproscan [118], release 5.16–55 (data package 55), with Hamap (201511.02), ProDom (2006.1), PIRSF (3.01), PANTHER (10.0), Pfam (28.0), SMART (6.2), Gene3D (3.5.0), Coils (2.2.1), ProSiteProfiles (20.113), TIGRFAM (15.0), PRINTS (42.0), SUPERFAMILY (1.75), and ProSitePatterns (20.113).

The same protocol of InterProScan annotation was performed to annotate the predicted set of proteins from eleven additional fungi. The names indicated are from the NCBI Taxonomy Database; commonly used synonyms and/or NCBI genome assembly accession numbers are in brackets: *Arthrobotrys oligospora* (ADOT00000000.1) [8], *Fusarium graminearum* (*Gibberella zeae*; AACM00000000.2) [119], *Fusarium oxysporum* (AAXH00000000.1) [120], *Hirsutella minnesotensis* (JPUM00000000.1) [9], *Metarhizium acridum* (ADNI00000000.1) [40], *Metarhizium anisopliae* (AZNF00000000.1) [40], *Monacrosporium haptotylum* (*Dactylellina haptotyla*; AQGS00000000.1) [5], *Ophiocordyceps sinensis* (ANOV00000000.1) [41], *Pochonia chlamydosporia* (*Metacordyceps chlamydospora*; AOSW00000000.1) [98], *Tolypocladium ophioglossoides* (*Elaphocordyceps ophioglossoides*; LFRF00000000.1) [64] and *Trichoderma reesei* (*Hypocrea jecorina*; AAIL00000000.2) [121].

SignalP v4.1 [122], TargetP v1.1 [123] and Tmhmm v2.0 [124] were used to predict respectively signal peptide, target peptide and transmembrane domains. A Blast analysis (BLASTP with e value $< 10^{-5}$) versus PHI-base proteins [65] was performed to associate *D. coniospora* genes to experimentally verified pathogenicity, virulence and effector genes from fungal, oomycete and bacterial pathogens.

Phylogenetic analysis

We refined the set of BUSCO-defined orthologues by restricting it to proteins that did not differ by more than 10% in total length across all 12 fungal species. This left us with a set of 97 high-confidence orthologous proteins present in all species. The respective sequences were concatenated were aligned using MAFFT [47] and phylogenetic distances calculated using the maximum likelihood-based method implemented within PhyML [48]. Altering the order of concatenation had no influence on the calculated phylogenetic distances. These analyses were performed within the Mobyle Web environment [125] at http://mobyle.pasteur.fr and the output plotted using the tree drawing engine implemented in the ETE toolkit [126].

CAZy analysis

Each *D. coniospora* protein model was compared using BLASTP [127] to proteins listed in the CAZy database (www.cazy.org; [128]). Because the e-value depends on the length of the aligned segment (for instance a 30% sequence identity results in widely different e-values, from non-significant to highly significant, if the two aligned proteins are 40, 100, 250 or 500 residues

in length), CAZy family assignments rather included examination of sequence conservation (percentage identity over CAZy domain length). Proteins that gave more than 50% identity over the entire domain length of an entry in CAZy were directly assigned to the same family. Proteins with less than 50% identify to a protein in CAZy were all manually inspected and conserved features such as catalytic residues were searched. The variable modular structure of CAZymes was integrated by performing alignments with isolated functional domains [129]. The same methods were used for all fungi that were compared to *D. coniospora*. For clustering of protein families and domains, we used "One minus Pearson correlation" distance matrices within GENE-E (www.broadinstitute.org/cancer/software/GENE-E/).

Secretome and comparative PFAM analysis

Secretome analysis was carried out as previously described [130] by combining predictions from SignalP v4.1 [122], WolfPSORT [131,132] and NucPred [133]. A WolfPSORT search using mature secreted proteins and model 'ANIMAL' was used to determine probable target protein localization in host. For sequence-based clustering, a database consisting of predicted proteomes of *D. coniospora*, *H. minnesotensis*, *P. chlamidosporia*, *M. haptotylum*, *A. oligospora* and PHIbase v3.6 [65] entries was built. The result of a BLASTP search of this database against itself with an e-value cutoff of 10^{-30} was used as input for clustering with the MCL program in Biolayout Express 3D [134]. Some of these clusters (S11 Table) were projected onto the Circos plot (Fig 1, in red, purple and green, respectively): OG5_126718 as putative nonribosomal peptide synthetases; OG5_127207 (serine carboxypeptidase S28), OG5_138644 (deuterolysin metalloprotease (M35) family), OG5_137388 (PA domain; subtilase family), OG5_128249 (subtilase family), OG5_149879 (serine carboxypeptidase), together with orthoMCL-defined paralogGroup 11 (subtilisin-like serine protease; S6 Table) as diverse proteases; paralogGroups 1, 4, 9 and 17 as enterotoxin-like proteins. Clusters that did not contain predicted secreted proteins from any of the 5 fungal species were discarded. Correspondence analysis was performed using the FactomineR package in R.

TargetP v1.1 [123] was used to determine probable protein localization, using mature proteins for secreted proteins or full-length sequences otherwise. A consensus predicted localization was derived using the following rules: TargetP predictions with reliability $< = 3$ only were considered; NLS were considered if predicted by NLStradamus and with a NucPred score <0.6; PredGPI predictions were considered if probability $> = 90$. PFAM domains were identified through a search against PFAM 28.0 database using gathering thresholds.

Gene expression analysis

For samples from mycelia and spores, Illumina paired-end (2 x 90 bp) RNAseq reads were aligned to the genome assembly using STAR [135]. Reads were assigned to the 8733 gene models using the htseq-count script within HTSeq [136]. To establish lists of differentially regulated genes, we used previous described methods [74] and retained the genes that were commonly defined by both. For samples from *C. elegans* infected with *D. coniospora*, the unaligned reads (640069 and 237794 reads from 5 and 12 h samples, respectively; kindly provided by LaDeana Hillier) from a previous RNAseq analysis [74] were aligned and assigned to gene models as above.

Protoplast preparation and fungal transformatn

N2 worms at the L4 stage were infected with fungal spores as described [114] on NGMY plates spread with the *E. coli* strain OP50 and incubated for 24 h, then transferred into NGMY liquid medium and cultured for another 30 h. Mycelia were collected by filtration as above and

protoplasts prepared and transformed using polyethylene glycol (PEG)/$CaCl_2$-mediated DNA uptake as described [137], using expression vectors (see S1 Methods) containing a hygromycin selection marker, derived from pPK2*hphgfp* [78], a kind gift from Martijn Rep (Swammerdam Institute for Life Sciences, Amsterdam). Transformants were selected for antibiotic resistance on medium containing 15 μg/ml hygromycin and screened for fluorescence.

Spore protein lysates and pull-down assays

Worms were infected on NGM plates with spores from SapA::DsRed expressing fungus. After 15 to 30 day culture at 25°C, spores were harvested in 50 mM NaCl as previously described [114]. They were extensively washed in cold 50 mM NaCl, pelleted and resuspended in an equal volume of lysis buffer (50mM Tris-Cl, pH 7.5, 100mM NaCl, 3 mM MgCl2, 0.5% Triton X-100, protease inhibitors (Complete, Roche), 5% glycerol) and flash-frozen. They were then sonicated on a high setting for 7 minutes (30 sec on and 30 sec off; Bioruptor, Diagenode,) and vortexed for 5 min with acid-washed glass beads (Sigma). The supernatant from a high-speed centrifugation was used as a whole protein extract for pull-down assays. For this, 800 μg of protein extracts were incubated with 20 μg of purified His-tagged SPP-1, SPP-5, or SPP-12 (the generous gift of M. Leippe, Kiel university), for 2 h. Preformed complexes were then immunoprecipitated with anti-dsRed/RFP agarose beads (Chromotek, RFP-Trap), at 4°C, overnight. As a binding control, preformed complexes were incubated under identical conditions with blocked matrix beads. After incubation, beads were washed three times in lysis buffer and three times in wash buffer (25 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% NP-40, 1 mM $MgCl_2$). After a final wash (1 mM Tris-Cl, 150 mM NaCl and 1 mM $MgCl_2$) beads were eluted in SDS-PAGE loading buffer. Samples were then resolved on a 4–12% gradient denaturating PAGE gel and transferred onto a membrane. The blot was probed with anti-His antibody (Upstate, H8 clone) or anti-dsRed/RFP antibody (Rockland, Inc). Blots were visualized using enhanced chemiluminescence (SuperSignal West Pico, Pierce).

Nucleotide sequence accession numbers

MIseq reads used for *de novo* genome assembly and Hiseq reads used for RNAseq *de novo* have been deposited at SRA under accession numbers SRX883538 and SRX969055, respectively. The Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JYHR00000000. The version described in this paper is version JYHR01000000.

Supporting Information

S1 Table. Descriptive statistics of the different assemblies before and after SSPACE scaffolding. Correspondence of the 9 optical maps (named as chromosomes) to the scaffolds from Velvet (kmer = 63), SOAPdenovo (kmer = 63), Spades (kmer = 127), and ABySS (kmer = 64 and 96), before and after scaffolding with SSPACE. Two values are provided, which correspond to the number of scaffolds localized on each optical map and the total coverage in Mb. The total number of mis-assemblies observed within each analysis is also indicated. ND: not determined.
(XLSX)

S2 Table. Comparative analysis of transposable elements in the genome of *D. coniospora* and of 11 other fungal species. The sources for the genome sequences used in this analysis are given in the Materials and Methods.
(XLSX)

S3 Table. Comparative analysis of tRNAs in the genome of *D. coniospora* and of 11 other fungal species.
(XLSX)

S4 Table. Species and sequences used for phylogenetic analysis.
(XLSX)

S5 Table. InterproScan analysis of the 8733 proteins predicted from the *D. coniospora* genome. The first sheet presents the results of an InterproScan analysis of the predicted proteome, complemented with dedicated BLASTP searches against UniprotKB and PHI-base databases. The second gives the sequences of 31 predicted proteins longer than 1000 amino acids, but without any InterproScan annotation. The third sheet gives a summary of the BLASTP analysis against the NCBI non-redundant set of proteins for the 31 sequences on the previous sheet. The results for the predicted proteins are split into 3 groups: with a homologue (only top 10 hits shown), with only remote similarity to other proteins (all hits shown), and with no similarity to any protein in the current NCBI non-redundant protein sequence database.
(XLSX)

S6 Table. OrthoMCL analysis of the 8733 proteins predicted from the *D. coniospora* genome. The first sheet lists the species present in OrthoMCL, with the frequency of top hits per species. The second sheet indicates the top OrthoMCL hit for the 7092 *D. coniospora* proteins assigned to an OrthoMCL cluster. The third sheet gives the number of predicted proteins in each OrthoMCL cluster, with corresponding PFAM domain annotation (from S5 Table) for those with more than 3 members. The fourth sheet lists the members of the clusters determined by OrthoMCL to be specific to *D. coniospora* (i.e. not represented in the 150 species in the first sheet). The fifth sheet gives the CLUSTAL Omega alignments for the 15 predicted proteins in the OrthoMCL paralogous group 1 (from fourth sheet).
(XLSX)

S7 Table. Comparative analysis of the occurrence of PFAM domains in the 8733 proteins predicted from the *D. coniospora* genome and in the predicted proteins from 11 other fungal species. Only PFAM domains present in at least one species are shown. The first sheet shows all the data. The second and third, extracted from the first, shows PFAM domains absent and present in *D. coniospora*, respectively.
(XLSX)

S8 Table. Detailed analysis of *D. coniospora* CAZy proteins and comparative analysis with the predicted CAZy proteins from 10 other fungal species. The first sheet indicates the classification of the predicted 240 CAZy proteins in *D. coniospora*, with notes from manual annotation. The following sheets show the occurrence of different CAZy family proteins in *D. coniospora* and 10 other fungi. Only CAZy families present in at least one species are shown.
(XLSX)

S9 Table. Predicted *D. coniospora* proteins involved in the production of secondary metabolites. The first sheet lists the proteins identified by SMURF analysis as 'backbone'; the second lists the proteins that putatively form part of a functional cluster.
(XLSX)

S10 Table. Analysis of hits in the PHI-base database for the 8,733 proteins predicted from the *D. coniospora* genome. The first sheet gives the number of times a given PHI-base entry was returned as the top hit for a *D. coniospora* protein (indicated in the column 'Occurrence'; data derived from the column 'PHIBASE' in S5 Table). The second sheet lists the PHI-base

entries for those hit by 5 or more predicted *D. coniospora* proteins. Certain indicated functional categories are highlighted in colour. The third sheet is an extract of the InterproScan data in S5 Table for all the *D. coniospora* proteins listed in the second sheet, and the following 7 sheets are extracts for the proteins corresponding to the PHI-base entries hit at least 10 times.
(XLSX)

S11 Table. Details of the secretome analysis for the 670 *D. coniospora* predicted with high confidence to be secreted. The first sheet lists the *D. coniospora* proteins predicted to be secreted, together with the scores from the analysis programs, and annotations taken from S5 Table. The second sheet gives the data sources for the comparative analysis. The last sheet shows the results of sequence based clustering of proteins from five nematophagous fungal species. Only clusters including at least one predicted secreted protein are considered.
(XLSX)

S12 Table. RNAseq analysis of transcripts from spore and mycelial samples. The first sheet gives the read counts for each gene from the 2 different samples. The 50 most highly expressed genes for each sample are highlighted. In the next 2 sheets they are listed together with annotations taken from S5 Table. The fourth and the fifth sheets list the genes assigned to the over-represented category in spores and in mycelia, respectively, together with annotations taken from S5 Table. On the fourth sheet, neighbouring genes are highlighted in yellow. The sixth and seventh list their respective constituent PFAM domains, with a score that reflects the confidence of the assignment. Seven domains are found in both lists (PF00005, PF00083, PF00172, PF00501, PF05730, PF07690, PF13193).
(XLSX)

S13 Table. RNAseq analysis of *D. coniospora* transcripts from samples extracted from infected *C. elegans*. The first sheet gives the raw read counts for genes covered by at least one RNAseq read from the 2 samples, listed by name and total read counts. The second sheet lists the genes covered by at least 3 reads, gives the read counts for the samples from spores and mycelia (data from S12 Table; genes in the top 15^{th} percentile highlighted) and includes different functional annotations (from S5 Table).
(XLSX)

S1 Fig. RADAR analysis [54] reveals the repeated structure in the sequence of g8068.t1, a 1045 a.a. protein from OrthoMCL-defined paralogous group 2 (see S6 Table).
(PDF)

S2 Fig. Photomicrographs of mycelial (A, B, C) or spore (D) preparations of *D. coniospora*, taken shortly before processing for RNA extraction. (A) Under the conditions of liquid culture used, *D. coniospora* forms compact balls of up to several mm in diameter. (B, C) At a higher magnification, it can be seen that the mycelia are devoid of spores and the fusion of hyphae can be clearly observed (white arrows). (D) While the majority of spores have started to germinate (red arrows), some have not (white arrows). A smaller proportion is not mature, lacking the adhesive bud (yellow arrows). Scale bars (white) in C and D, 10 μm.
(PDF)

S3 Fig. PCR-based verification of the insertion of a hygromycin-resistance expression cassette into the *Dso* locus. The top part of the figure shows the position of PCR primers relative to the genomic and recombinant DNA sequences. The 2 tables indicate the expected sizes and occurrences of PCR amplicons. The bottom part of the figure shows that the expected bands are obtained from the wild-type (WT) and knocked-in strain (Dso).
(PDF)

S4 Fig. (A) PCR products from reverse-transcribed mRNA corresponding to the *D. coniospora* saposin A-domain protein-encoding gene g3895.t1 (SapA) and the actin gene g2551.t1 (Actin) from mycelia and at the indicated times post-infection (p.i.). The size markers in the outside lanes are, from top to bottom, 300, 200 and 100 bp. (B) Clustal multiple alignment of infection-induced saposin proteins (in bold) and those used for to assay for a possible interaction between a host saposin and the fungal SapA protein.
(PDF)

S1 Text. Contains comments on the quality of the genome sequence and gene annotation, as well as a description of the isolation history of ATCC 96282 and its derivatives.
(PDF)

S1 Methods. Contains details of media, plasmid constructions and primer sequences.
(DOCX)

Acknowledgments

We thank Jérôme Gouzy for his input on an initial round of genome annotation and Aude Barani, Jérôme Belougne, Carole Couillault, Stéphane Coulon, Jan Dijksterhuis, Laurent Hannouche, LaDeana Hillier, Matthias Leippe, Luis Lopez-Llorca, Hans-Börge Jansson, Martijn Rep and particularly Nathalie Pujol for their contributions, advice and/or the generous gift of reagents. We also thank the anonymous reviewers for their constructive criticisms.

Author Contributions

Conceived and designed the experiments: KL JP ER VB SR PB JJE. Performed the experiments: KL LDH NT MJA JP GM SR JJE. Analyzed the data: KL BH NT ER SR PB JJE. Contributed reagents/materials/analysis tools: KL NT SR. Wrote the paper: KL NT BH ER PB JJE.

References

1. Barron GL. Nematophagous destroying fungi. Guelph: Lancester press; 1977. 1–140 p.
2. Linford MB. Stimulated activity of natural enemies of nematodes. Science. 1937; 85:123–4. PMID: 17754209
3. Ahren D, Tholander M, Fekete C, Rajashekar B, Friman E, Johansson T, et al. Comparison of gene expression in trap cells and vegetative hyphae of the nematophagous fungus *Monacrosporium haptotylum*. Microbiology. 2005; 151(Pt 3):789–803. PMID: 15758225
4. Fekete C, Tholander M, Rajashekar B, Ahren D, Friman E, Johansson T, et al. Paralysis of nematodes: shifts in the transcriptome of the nematode-trapping fungus *Monacrosporium haptotylum* during infection of *Caenorhabditis elegans*. Environ Microbiol. 2008; 10(2):364–75. PMID: 18028414
5. Meerupati T, Andersson KM, Friman E, Kumar D, Tunlid A, Ahren D. Genomic mechanisms accounting for the adaptation to parasitism in nematode-trapping fungi. PLoS Genet. 2013; 9(11):e1003909. Epub 2013/11/19. doi: 10.1371/journal.pgen.1003909 PMID: 24244185
6. Andersson KM, Kumar D, Bentzer J, Friman E, Ahren D, Tunlid A. Interspecific and host-related gene expression patterns in nematode-trapping fungi. Bmc Genomics. 2014; 15:968. Epub 2014/11/12. doi: 10.1186/1471-2164-15-968 PMID: 25384908
7. Liu K, Zhang W, Lai Y, Xiang M, Wang X, Zhang X, et al. *Drechslerella stenobrocha* genome illustrates the mechanism of constricting rings and the origin of nematode predation in fungi. Bmc Genomics. 2014; 15:114. Epub 2014/02/11. doi: 10.1186/1471-2164-15-114 PMID: 24507587
8. Yang J, Wang L, Ji X, Feng Y, Li X, Zou C, et al. Genomic and Proteomic Analyses of the Fungus *Arthrobotrys oligospora* Provide Insights into Nematode-Trap Formation. PLoS Pathog. 2011; 7(9): e1002179. Epub 2011/09/13. doi: 10.1371/journal.ppat.1002179 PMID: 21909256
9. Lai Y, Liu K, Zhang X, Li K, Wang N, Shu C, et al. Comparative genomics and transcriptomics analyses reveal divergent lifestyle features of nematode endoparasitic fungus *Hirsutella minnesotensis*. Genome Biol Evol. 2014; 6(11):3077–93. Epub 2014/11/02. doi: 10.1093/gbe/evu241 PMID: 25359922

10. Dijksterhuis J, Veenhuis M, Harder W. Ultrastructural study of adhesion and initial stages of infection of the nematode by conidia of *Drechmeria coniospora*. Mycological research. 1990; 94(1):1–8.
11. Rouger V, Bordet G, Couillault C, Monneret S, Mailfert S, Ewbank JJ, et al. Independent Synchronized Control and Visualization of Interactions between Living Cells and Organisms. Biophysical journal. 2014; 106(10):2096–104. doi: 10.1016/j.bpj.2014.03.044 PMID: 24853738
12. Gernandt DS, Stone JK. Phylogenetic analysis of nuclear ribosomal DNA place the nematode parasite, *Drechmeria coniospora*, in Clavicipitaceae. Mycologia. 1999; 91(6):993–1000.
13. Quandt CA, Kepler RM, Gams W, Araujo JP, Ban S, Evans HC, et al. Phylogenetic-based nomenclatural proposals for *Ophiocordycipitaceae* (*Hypocreales*) with new combinations in *Tolypocladium*. IMA Fungus. 2014; 5(1):121–34. doi: 10.5598/imafungus.2014.05.01.12 PMID: 25083412
14. Jansson HB, Jeyaprakash A, Zuckerman BM. Differential adhesion and infection of nematodes by the endoparasitic fungus *Meria coniospora* (*Deuteromycetes*). Appl Envir Microbiol. 1985; 49:552–5.
15. Coles GC, Dicklow MB, Zuckerman BM. Protein changes associated with the infection of the nematode *Caenorhabditis elegans* by the nematophagous fungus *Drechmeria coniospora*. Int J Parasitol. 1989; 19:733–6.
16. Fuchs BB, Mylonakis E. Using non-mammalian hosts to study fungal virulence and host defense. Curr Opin Microbiol. 2006; 9(4):346–51. PMID: 16814595
17. Ewbank JJ, Zugasti O. *C. elegans*: model host and tool for antimicrobial drug discovery. Dis Model Mech. 2011; 4(3):300–4. Epub 2011/04/21. doi: 10.1242/dmm.006684 PMID: 21504910
18. Clark LC, Hodgkin J. Commensals, probiotics and pathogens in the *Caenorhabditis elegans* model. Cellular microbiology. 2014; 16(1):27–38. Epub 2013/10/31. doi: 10.1111/cmi.12234 PMID: 24168639
19. Cohen LB, Troemel ER. Microbial pathogenesis and host defense in the nematode *C. elegans*. Curr Opin Microbiol. 2015; 23C:94–101. Epub 2014/12/03.
20. Labed S, Pujol N. *Caenorhabditis elegans* Antifungal Defense Mechanisms. The Journal of Invasive Fungal Infection. 2011; 5(4):110–7.
21. Ewbank JJ, Pujol N. Local and long-range activation of innate immunity by infection and damage in *C. elegans*. Curr Opin Immunol. 2016; 38:1–7. doi: 10.1016/j.coi.2015.09.005 PMID: 26517153
22. Kim DH, Ewbank JJ. Signaling in the Immune Response. 2015 Dec 22. In: WormBook [Internet]. http://www.wormbook.org; [1–51]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26694508.
23. Felix MA, Duveau F. Population dynamics and habitat sharing of natural populations of *Caenorhabditis elegans* and *C. briggsae*. BMC Biol. 2012; 10(1):59. Epub 2012/06/27.
24. Zugasti O, Bose N, Squiban B, Belougne J, Kurz CL, Schroeder FC, et al. Activation of a G protein-coupled receptor by its endogenous ligand triggers the innate immune response of *Caenorhabditis elegans*. Nat Immunol. 2014; 15(9):833–8. Epub 2014/08/05. doi: 10.1038/ni.2957 PMID: 25086774
25. Squiban B, Belougne J, Ewbank J, Zugasti O. Quantitative and automated high-throughput genome-wide RNAi screens in *C. elegans*. J Vis Exp. 2012; 60:e3448. Epub 2012/03/08.
26. Labed SA, Omi S, Gut M, Ewbank JJ, Pujol N. The pseudokinase NIPI-4 is a novel regulator of antimicrobial peptide gene expression. PLoS One. 2012; 7(3):e33887. Epub 2012/04/04. doi: 10.1371/journal.pone.0033887 PMID: 22470487
27. Dierking K, Polanowska J, Omi S, Engelmann I, Gut M, Lembo F, et al. Unusual regulation of a STAT protein by an SLC6 family transporter in *C. elegans* epidermal innate immunity. Cell Host Microbe. 2011; 9(5):425–35. Epub 2011/05/18. doi: 10.1016/j.chom.2011.04.011 PMID: 21575913
28. Zugasti O, Ewbank JJ. Neuroimmune regulation of antimicrobial peptide expression by a noncanonical TGF-beta signaling pathway in *Caenorhabditis elegans* epidermis. Nat Immunol. 2009; 10(3):249–56. doi: 10.1038/ni.1700 PMID: 19198592
29. Nomura K, Debroy S, Lee YH, Pumplin N, Jones J, He SY. A bacterial virulence protein suppresses host innate immunity to cause plant disease. Science. 2006; 313(5784):220–3. Epub 2006/07/15. PMID: 16840699
30. Elde NC, Malik HS. The evolutionary conundrum of pathogen mimicry. Nat Rev Microbiol. 2009; 7 (11):787–97. Epub 2009/10/07. doi: 10.1038/nrmicro2222 PMID: 19806153
31. Kepp O, Senovilla L, Galluzzi L, Panaretakis T, Tesniere A, Schlemmer F, et al. Viral subversion of immunogenic cell death. Cell Cycle. 2009; 8(6):860–9. Epub 2009/02/18. PMID: 19221507
32. Shames SR, Auweter SD, Finlay BB. Co-evolution and exploitation of host cell signaling pathways by bacterial pathogens. Int J Biochem Cell Biol. 2009; 41(2):380–9. Epub 2008/09/09. doi: 10.1016/j.biocel.2008.08.013 PMID: 18775503

33. Pujol N, Zugasti O, Wong D, Couillault C, Kurz CL, Schulenburg H, et al. Anti-fungal innate immunity in *C. elegans* is enhanced by evolutionary diversification of antimicrobial peptides. PLoS Pathog. 2008; 4(7):e1000105. Epub 2008/07/19. doi: 10.1371/journal.ppat.1000105 PMID: 18636113

34. Zerbino DR, Birney E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 2008; 18:821–9. doi: 10.1101/gr.074492.107 PMID: 18349386

35. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of computational biology: a journal of computational molecular cell biology. 2012; 19(5):455–77.

36. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. GigaScience. 2012; 1(1):18-. doi: 10.1186/2047-217X-1-18 PMID: 23587118

37. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol Iß. ABySS: a parallel assembler for short read sequence data. Genome Res. 2009; 19(6):1117–23. doi: 10.1101/gr.089532.108 PMID: 19251739

38. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics. 2011; 27(4):578–9. Epub 2010/12/15. doi: 10.1093/bioinformatics/btq683 PMID: 21149342

39. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. Genome Res. 2009; 19(9):1639–45. Epub 2009/06/23. doi: 10.1101/gr.092759.109 PMID: 19541911

40. Gao Q, Jin K, Ying SH, Zhang Y, Xiao G, Shang Y, et al. Genome sequencing and comparative transcriptomics of the model entomopathogenic fungi *Metarhizium anisopliae* and *M. acridum*. PLoS Genet. 2011; 7(1):e1001264. Epub 2011/01/22. doi: 10.1371/journal.pgen.1001264 PMID: 21253567

41. Hu X, Zhang Y, Xiao G, Zheng P, Xia Y, Zhang X, et al. Genome survey uncovers the secrets of sex and lifestyle in caterpillar fungus Chinese Science Bulletin. 2013; 58(23):2846–54.

42. Price AL, Jones NC, Pevzner PA. De novo identification of repeat families in large genomes. Bioinformatics (Oxford, England). 2005; 21 Suppl 1:i351–8.

43. Lowe TM, Eddy SR. tRNAscan-SE: A Program for Improved Detection of Transfer RNA Genes in Genomic Sequence. Nucleic Acids Res. 1997; 25(5):0955–964.

44. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990; 215(3):403–10. Epub 1990/10/05. PMID: 2231712

45. Stanke M, Diekhans M, Baertsch R, Haussler D. Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. Bioinformatics. 2008; 24(5):637–44. doi: 10.1093/bioinformatics/btn013 PMID: 18218656

46. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015; 31(19):3210–2. doi: 10.1093/bioinformatics/btv351 PMID: 26059717

47. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013; 30(4):772–80. doi: 10.1093/molbev/mst010 PMID: 23329690

48. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol. 2003; 52(5):696–704. PMID: 14530136

49. Lin R, Liu C, Shen B, Bai M, Ling J, Chen G, et al. Analysis of the complete mitochondrial genome of *Pochonia chlamydosporia* suggests a close relationship to the invertebrate-pathogenic fungi in Hypocreales. BMC Microbiol. 2015; 15(1):5. Epub 2015/02/01.

50. Zdobnov EM, Apweiler R. InterProScan—an integration platform for the signature-recognition methods in InterPro. Bioinformatics. 2001; 17(9):847–8. PMID: 11590104

51. Lipman DJ, Souvorov A, Koonin EV, Panchenko AR, Tatusova TA. The relationship of protein conservation and sequence length. BMC Evol Biol. 2002; 2:20. PMID: 12410938

52. Li L, Stoeckert CJ, Roos D. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. 2003; 13(9):2178–89. PMID: 12952885

53. Xiao G, Ying SH, Zheng P, Wang ZL, Zhang S, Xie XQ, et al. Genomic perspectives on the evolution of fungal entomopathogenicity in *Beauveria bassiana*. Sci Rep. 2012; 2:483. Epub 2012/07/05. doi: 10.1038/srep00483 PMID: 22761991

54. Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, et al. The EMBL-EBI bioinformatics web and programmatic tools framework. Nucleic Acids Res. 2015. Epub 2015/04/08.

55. Sonnhammer EL, Durbin R. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene. 1995; 167(1–2):GC1–10. PMID: 8566757

56. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, et al. The Pfam protein families database. Nucleic Acids Res. 2004; 32(Database issue):D138–41. PMID: 14681378

57. Habicht J, Woehle C, Gould SB. *Tetrahymena* Expresses More than a Hundred Proteins with Lipid-binding MORN Motifs that can Differ in their Subcellular Localisations. J Eukaryot Microbiol. 2015. Epub 2015/04/08.

58. Li J, Zhang KQ. Independent expansion of zincin metalloproteinases in Onygenales fungi may be associated with their pathogenicity. PLoS One. 2014; 9(2):e90225. Epub 2014/03/04. doi: 10.1371/journal.pone.0090225 PMID: 24587291

59. Li J, Yu L, Tian Y, Zhang KQ. Molecular evolution of the deuterolysin (M35) family genes in Coccidioides. PLoS One. 2012; 7(2):e31536. Epub 2012/03/01. doi: 10.1371/journal.pone.0031536 PMID: 22363666

60. Dyrka W, Lamacchia M, Durrens P, Kobe B, Daskalov A, Paoletti M, et al. Diversity and variability of NOD-like receptors in fungi. Genome Biol Evol. 2014; 6(12):3137–58. doi: 10.1093/gbe/evu251 PMID: 25398782

61. Fang H, Gough J. DcGO: database of domain-centric ontologies on functions, phenotypes, diseases and more. Nucleic Acids Res. 2013; 41(Database issue):D536–44. Epub 2012/11/20. doi: 10.1093/nar/gks1080 PMID: 23161684

62. Khaldi N, Seifuddin FT, Turner G, Haft D, Nierman WC, Wolfe KH, et al. SMURF: Genomic mapping of fungal secondary metabolite clusters. Fungal Genet Biol. 2010; 47(9):736–41. Epub 2010/06/18. doi: 10.1016/j.fgb.2010.06.003 PMID: 20554054

63. Scharf DH, Heinekamp T, Brakhage AA. Human and plant fungal pathogens: the role of secondary metabolites. PLoS Pathog. 2014; 10(1):e1003859. Epub 2014/02/06. doi: 10.1371/journal.ppat.1003859 PMID: 24497825

64. Quandt CA, Bushley KE, Spatafora JW. The genome of the truffle-parasite *Tolypocladium ophioglossoides* and the evolution of antifungal peptaibiotics. Bmc Genomics. 2015; 16(1):553.

65. Winnenburg R, Urban M, Beacham A, Baldwin TK, Holland S, Lindeberg M, et al. PHI-base update: additions to the pathogen host interaction database. Nucleic Acids Res. 2008; 36(Database issue): D572–6. Epub 2007/10/19. PMID: 17942425

66. Shen B, Xiao J, Dai L, Huang Y, Mao Z, Lin R, et al. Development of a high-efficiency gene knockout system for *Pochonia chlamydosporia*. Microbiol Res. 2015; 170:18–26. doi: 10.1016/j.micres.2014.10.001 PMID: 25458554

67. Hu X, Xiao G, Zheng P, Shang Y, Su Y, Zhang X, et al. Trajectory and genomic determinants of fungal-pathogen speciation and host adaptation. P Natl Acad Sci USA. 2014; 111(47):16796–801. Epub 2014/11/05.

68. Xue C, Hsueh YP, Heitman J. Magnificent seven: roles of G protein-coupled receptors in extracellular sensing in fungi. FEMS Microbiol Rev. 2008; 32(6):1010–32. Epub 2008/09/25. doi: 10.1111/j.1574-6976.2008.00131.x PMID: 18811658

69. Kulkarni RD, Kelkar HS, Dean RA. An eight-cysteine-containing CFEM domain unique to a group of fungal membrane proteins. Trends Biochem Sci. 2003; 28(3):118–21. Epub 2003/03/14. PMID: 12633989

70. Stringer MA, Dean RA, Sewall TC, Timberlake WE. *Rodletless*, a new *Aspergillus* developmental mutant induced by directed gene inactivation. Genes Dev. 1991; 5(7):1161–71. PMID: 2065971

71. Linder T, Gustafsson CM. Molecular phylogenetics of ascomycotal adhesins—a novel family of putative cell-surface adhesive proteins in fission yeasts. Fungal Genet Biol. 2008; 45(4):485–97. Epub 2007/09/18. PMID: 17870620

72. Ohtaki S, Maeda H, Takahashi T, Yamagata Y, Hasegawa F, Gomi K, et al. Novel hydrophobic surface binding protein, HsbA, produced by *Aspergillus oryzae*. Applied and environmental microbiology. 2006; 72(4):2407–13. Epub 2006/04/07. PMID: 16597938

73. Yang J, Zhao X, Liang L, Xia Z, Lei L, Niu X, et al. Overexpression of a cuticle-degrading protease Ver112 increases the nematicidal activity of *Paecilomyces lilacinus*. Appl Microbiol Biotechnol. 2011; 89(6):1895–903. Epub 2010/11/27. doi: 10.1007/s00253-010-3012-6 PMID: 21110018

74. Engelmann I, Griffon A, Tichit L, Montanana-Sanchis F, Wang G, Reinke V, et al. A comprehensive analysis of gene expression changes provoked by bacterial and fungal infection in *C. elegans*. PLoS One. 2011; 6(5):e19055. Epub 2011/05/24. doi: 10.1371/journal.pone.0019055 PMID: 21602919

75. Teixeira PJ, Thomazella DP, Vidal RO, do Prado PF, Reis O, Baroni RM, et al. The fungal pathogen *Moniliophthora perniciosa* has genes similar to plant PR-1 that are highly expressed during its interaction with cacao. PLoS One. 2012; 7(9):e45929. doi: 10.1371/journal.pone.0045929 PMID: 23029323

76. Sakamoto Y, Nakade K, Konno N. Endo-beta-1,3-glucanase GLU1, from the fruiting body of *Lentinula edodes*, belongs to a new glycoside hydrolase family. Appl Environ Microbiol. 2011; 77(23):8350–4. doi: 10.1128/AEM.05581-11 PMID: 21965406

77. Punt PJ, van den Hondel CA. Transformation of filamentous fungi based on hygromycin B and phleomycin resistance markers. Methods Enzymol. 1992; 216:447–57. Epub 1992/01/01. PMID: 1479914

78. Michielse CB, van Wijk R, Reijnen L, Cornelissen BJ, Rep M. Insight into the molecular requirements for pathogenicity of *Fusarium oxysporum* f. sp. *lycopersici* through large-scale insertional mutagenesis. Genome biology. 2009; 10(1):R4. Epub 2009/01/13. doi: 10.1186/gb-2009-10-1-r4 PMID: 19134172

79. Fleissner A, Sarkar S, Jacobson DJ, Roca MG, Read ND, Glass NL. The *so* locus is required for vegetative cell fusion and postfertilization events in *Neurospora crassa*. Eukaryot Cell. 2005; 4(5):920–30. Epub 2005/05/10. PMID: 15879526

80. Prados Rosales RC, Di Pietro A. Vegetative hyphal fusion is not essential for plant infection by *Fusarium oxysporum*. Eukaryot Cell. 2008; 7(1):162–71. Epub 2007/11/28. PMID: 18039941

81. Duverger Y, Belougne J, Scaglione S, Brandli D, Beclin C, Ewbank JJ. A semi-automated high-throughput approach to the generation of transposon insertion mutants in the nematode *Caenorhabditis elegans*. Nucleic Acids Res. 2007; 35(2):e11. PMID: 17164286

82. Banyai L, Patthy L. Amoebapore homologs of *Caenorhabditis elegans*. Biochim Biophys Acta. 1998; 1429(1):259–64. PMID: 9920402

83. Roeder T, Stanisak M, Gelhaus C, Bruchhaus I, Grotzinger J, Leippe M. Caenopores are antimicrobial peptides in the nematode *Caenorhabditis elegans* instrumental in nutrition and immunity. Dev Comp Immunol. 2010; 34(2):203–9. Epub 2009/10/13. doi: 10.1016/j.dci.2009.09.010 PMID: 19818806

84. Mysliwy J, Dingley AJ, Stanisak M, Jung S, Lorenzen I, Roeder T, et al. Caenopore-5: the three-dimensional structure of an antimicrobial protein from *Caenorhabditis elegans*. Dev Comp Immunol. 2010; 34(3):323–30. doi: 10.1016/j.dci.2009.11.003 PMID: 19917307

85. Wong D, Bazopoulou D, Pujol N, Tavernarakis N, Ewbank JJ. Genome-wide investigation reveals pathogen-specific and shared signatures in the response of *Caenorhabditis elegans* to infection. Genome Biol. 2007; 8(9):R194. PMID: 17875205

86. Hoeckendorf A, Leippe M. SPP-3, a saposin-like protein of *Caenorhabditis elegans*, displays antimicrobial and pore-forming activity and is located in the intestine and in one head neuron. Dev Comp Immunol. 2012; 38(1):181–6. doi: 10.1016/j.dci.2012.05.007 PMID: 22677064

87. Hoeckendorf A, Stanisak M, Leippe M. The saposin-like protein SPP-12 is an antimicrobial polypeptide in the pharyngeal neurons of *Caenorhabditis elegans* and participates in defence against a natural bacterial pathogen. Biochem J. 2012; 445(2):205–12. Epub 2012/04/24. doi: 10.1042/BJ20112102 PMID: 22519640

88. Drechsler C. Some hyphomycetes parasitic on free-living terricolous nematodes. Phytopathology. 1941; 31:773–802.

89. Jansson HB, Jeyaprakash A, Zuckerman BM. Control of Root-Knot Nematodes on Tomato by the Endoparasitic Fungus *Meria coniospora*. J Nematol. 1985; 17(3):327–9. PMID: 19294101

90. Meyer SL, Huettel RN, Sayre RM. Isolation of Fungi from *Heterodera glycines* and *in vitro* Bioassays for Their Antagonism to Eggs. J Nematol. 1990; 22(4):532–7. PMID: 19287754

91. Poinar GO, Jansson HB. Susceptibility of *Neoaplectana* spp. and *Heterorhabditis heliothidis* to the Endoparasitic Fungus *Drechmeria coniospora*. J Nematol. 1986; 18(2):225–9. Epub 1986/04/01. PMID: 19294171

92. Dijksterhuis J, Veenhuis M, Harder W. Conidia of the nematophagous fungus *Drechmeria coniospora* adhere to but barely infect *Acrobeloides buetschilii*. FEMS Microbiology Letters. 1993; 113(2):183–8.

93. Jansson HB. Adhesion to Nematodes of Conidia from the Nematophagous Fungus *Drechmeria coniospora*. Journal of General Microbiology. 1993; 139:1899–906.

94. van den Boogert PH, Dijksterhuis J, Velvis H, Veenhuis M. Adhesive knob formation by conidia of the nematophagous fungus *Drechmeria coniospora*. Antonie Van Leeuwenhoek. 1992; 61(3):221–9. PMID: 1519917

95. Sun J, Park SY, Kang S, Liu X, Qiu J, Xiang M. Development of a transformation system for *Hirsutella spp*. and visualization of the mode of nematode infection by GFP-labeled *H. minnesotensis*. Sci Rep. 2015; 5:10477. doi: 10.1038/srep10477 PMID: 26190283

96. Rosso LC, Finetti-Sialer MM, Hirsch PR, Ciancio A, Kerry BR, Clark IM. Transcriptome analysis shows differential gene expression in the saprotrophic to parasitic transition of *Pochonia chlamydosporia*. Appl Microbiol Biotechnol. 2011; 90(6):1981–94. doi: 10.1007/s00253-011-3282-7 PMID: 21541788

97. Olivares CM, Lopez-Llorca LV. Fungal egg-parasites of plant-parasitic nematodes from Spanish soils. Rev Iberoam Micol. 2002; 19(2):104–10. PMID: 12828513

98. Larriba E, Jaime MD, Carbonell-Caballero J, Conesa A, Dopazo J, Nislow C, et al. Sequencing and functional analysis of the genome of a nematode egg-parasitic fungus, *Pochonia chlamydosporia*. Fungal Genet Biol. 2014; 65:69–80. Epub 2014/02/18. doi: 10.1016/j.fgb.2014.02.002 PMID: 24530791

99. Tahseen Q, Clark IM, Atkins SD, Hirsch PR, Kerry BR. Impact of the nematophagous fungus *Pochonia chlamydosporia* on nematode and microbial populations. Commun Agric Appl Biol Sci. 2005; 70 (1):81–6. PMID: 16363363

100. Mendoza De Gives PM, Davies KG, Clark SJ, Behnke JM. Predatory behaviour of trapping fungi against *srf* mutants of *Caenorhabditis elegans* and different plant and animal parasitic nematodes. Parasitology. 1999; 119(1):95–104.

101. Migunova VD, Byzov BA. Determinants of trophic modes of the nematophagous fungus *Arthrobotrys oligospora* interacting with bacterivorous nematode *Caenorhabditis elegans*. Pedobiologia. 2005; 49 (2):101–8.

102. Niu X-M, Zhang K-Q. *Arthrobotrys oligospora*: a model organism for understanding the interaction between fungi and nematodes. Mycology. 2011; 2(2):59–78.

103. van den Brink J, de Vries RP. Fungal enzyme sets for plant polysaccharide degradation. Appl Microbiol Biotechnol. 2011; 91(6):1477–92. Epub 2011/07/26. doi: 10.1007/s00253-011-3473-2 PMID: 21785931

104. Dijksterhuis J, Harder W, Wyss U, Veenhuis M. Colonization and digestion of nematodes by the endoparasitic nematophagous fungus *Drechmeria coniospora*. Mycological Research. 1991; 95:873–8.

105. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. Biotechnol Biofuels. 2013; 6(1):41. doi: 10.1186/1754-6834-6-41 PMID: 23514094

106. Fang W, Leger RJ St. *Mrt*, a gene unique to fungi, encodes an oligosaccharide transporter and facilitates rhizosphere competency in *Metarhizium robertsii*. Plant Physiol. 2010; 154(3):1549–57. Epub 2010/09/15. doi: 10.1104/pp.110.163014 PMID: 20837701

107. van der Biezen EA, Jones JD. The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. Curr Biol. 1998; 8(7):R226–7. PMID: 9545207

108. Fang W, Pava-ripoll M, Wang S, Leger R St. Protein kinase A regulates production of virulence determinants by the entomopathogenic fungus, *Metarhizium anisopliae*. Fungal Genet Biol. 2009; 46 (3):277–85. Epub 2009/01/07. doi: 10.1016/j.fgb.2008.12.001 PMID: 19124083

109. Lu J, Cao H, Zhang L, Huang P, Lin F. Systematic analysis of Zn_2Cys_6 transcription factors required for development and pathogenicity by high-throughput gene knockout in the rice blast fungus. PLoS Pathog. 2014; 10(10):e1004432. Epub 2014/10/10. doi: 10.1371/journal.ppat.1004432 PMID: 25299517

110. Dijksterhuis J, Sjollema KA, Veenhuis M, Harder W. Competitive interactions between two nematophagous fungi during infection and digestion of the nematode *Panagrellus redivivus*. Mycological Research. 1994; 98(12):1458–62.

111. Zou CG, Tao N, Liu WJ, Yang JK, Huang XW, Liu XY, et al. Regulation of subtilisin-like protease prC expression by nematode cuticle in the nematophagous fungus *Clonostachys rosea*. Environ Microbiol. 2010; 12(12):3243–52. doi: 10.1111/j.1462-2920.2010.02296.x PMID: 20636375

112. Muszewska A, Taylor JW, Szczesny P, Grynberg M. Independent subtilases expansions in fungi associated with animals. Mol Biol Evol. 2011; 28(12):3395–404. doi: 10.1093/molbev/msr176 PMID: 21727238

113. Baxter L, Tripathy S, Ishaque N, Boot N, Cabral A, Kemen E, et al. Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. Science. 2010; 330(6010):1549–51. Epub 2010/12/15. doi: 10.1126/science.1195203 PMID: 21148394

114. Powell JR, Ausubel FM. Models of *Caenorhabditis elegans* Infection by Bacterial and Fungal Pathogens. In: Ewbank J, Vivier E, editors. Methods Mol Biol. 415: Humana Press; 2008. p. 403–27.

115. Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem. 1987; 162(1):156–9. Epub 1987/04/01. PMID: 2440339

116. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011; 29(7):644–52. doi: 10.1038/nbt.1883 PMID: 21572440

117. Gardner PP, Daub J, Tate JG, Nawrocki EP, Kolbe DL, Lindgreen S, et al. Rfam: updates to the RNA families database. Nucleic Acids Res. 2009; 37(Database issue):D136–40. doi: 10.1093/nar/gkn766 PMID: 18953034

118. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, et al. InterProScan 5: Genome-scale protein function classification. Bioinformatics. 2014; 30:1236–40. doi: 10.1093/bioinformatics/btu031 PMID: 24451626

119. Cuomo CA, Guldener U, Xu JR, Trail F, Turgeon BG, Di Pietro A, et al. The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. Science. 2007; 317(5843):1400–2. Epub 2007/09/08. PMID: 17823352

120. Ma LJ, van der Does HC, Borkovich KA, Coleman JJ, Daboussi MJ, Di Pietro A, et al. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature. 2010; 464(7287):367–73. Epub 2010/03/20. doi: 10.1038/nature08850 PMID: 20237561

121. Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, et al. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). Nat Biotechnol. 2008; 26(5):553–60. Epub 2008/05/06. doi: 10.1038/nbt1403 PMID: 18454138

122. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature methods. 2011; 8(10):785–6. Epub 2011/10/01. doi: 10.1038/nmeth.1701 PMID: 21959131

123. Hawkins J, Boden M. Detecting and sorting targeting peptides with neural networks and support vector machines. J Bioinform Comput Biol. 2006; 4(1):1–18. Epub 2006/03/29. PMID: 16568539

124. Sonnhammer EL, von Heijne G, Krogh A. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc Int Conf Intell Syst Mol Biol. 1998; 6:175–82. Epub 1998/10/23. PMID: 9783223

125. Neron B, Menager H, Maufrais C, Joly N, Maupetit J, Letort S, et al. Mobyle: a new full web bioinformatics framework. Bioinformatics. 2009; 25(22):3005–11. doi: 10.1093/bioinformatics/btp493 PMID: 19689959

126. Huerta-Cepas J, Serra F, Bork P. ETE 3: Reconstruction, analysis and visualization of phylogenomic data. Mol Biol Evol. 2016.

127. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997; 25 (17):3389–402. Epub 1997/09/01. PMID: 9254694

128. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 2014; 42(Database issue):D490–5. Epub 2013/11/26. doi: 10.1093/nar/gkt1178 PMID: 24270786

129. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res. 2009; 37 (Database issue):D233–8. Epub 2008/10/08. doi: 10.1093/nar/gkn663 PMID: 18838391

130. Badet T, Peyraud R, Raffaele S. Common protein sequence signatures associate with *Sclerotinia borealis* lifestyle and secretion in fungal pathogens of the Sclerotiniaceae. Front Plant Sci. 2015; 6:776. doi: 10.3389/fpls.2015.00776 PMID: 26442085

131. Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, et al. WoLF PSORT: protein localization predictor. Nucleic Acids Res. 2007; 35(Web Server issue):W585–7. PMID: 17517783

132. Nguyen Ba AN, Pogoutse A, Provart N, Moses AM. NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction. BMC Bioinformatics. 2009; 10:202. Epub 2009/07/01. doi: 10.1186/1471-2105-10-202 PMID: 19563654

133. Brameier M, Krings A, MacCallum RM. NucPred—predicting nuclear localization of proteins. Bioinformatics. 2007; 23(9):1159–60. Epub 2007/03/03. PMID: 17332022

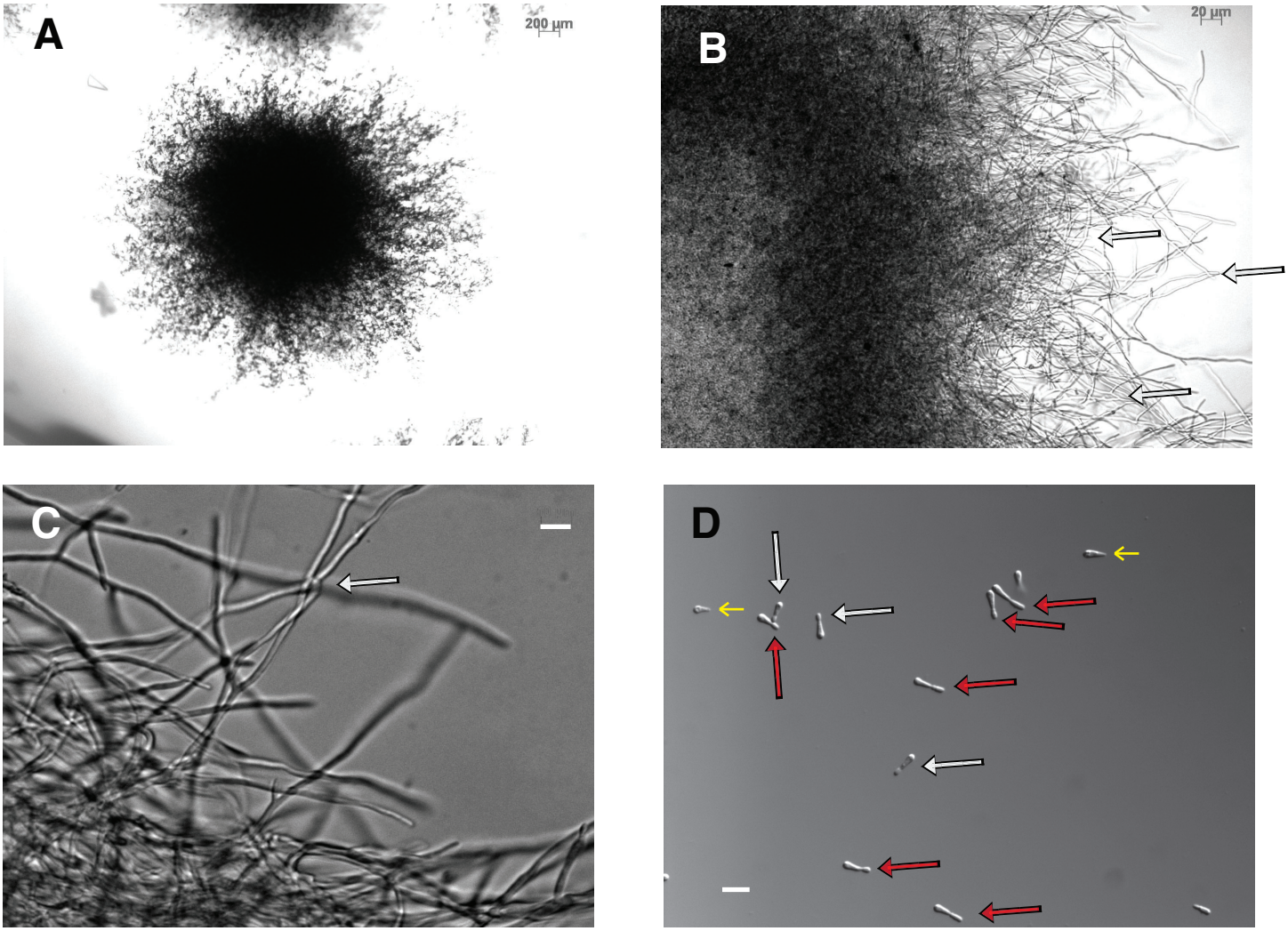
134. Theocharidis A, van Dongen S, Enright AJ, Freeman TC. Network visualization and analysis of gene expression data using BioLayout Express(3D). Nat Protoc. 2009; 4(10):1535–50. doi: 10.1038/nprot.2009.177 PMID: 19798086

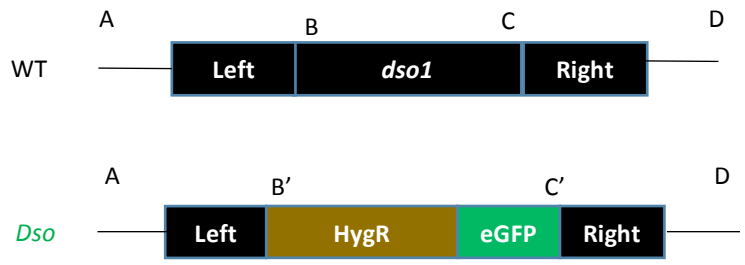
135. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013; 29(1):15–21. doi: 10.1093/bioinformatics/bts635 PMID: 23104886

136. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015; 31(2):166–9. doi: 10.1093/bioinformatics/btu638 PMID: 25260700

137. Turgeon BG, Condon B, Liu J, Zhang N. Protoplast transformation of filamentous fungi. Methods Mol Biol. 2010; 638:3–19. Epub 2010/03/20. doi: 10.1007/978-1-60761-611-5_1 PMID: 20238257

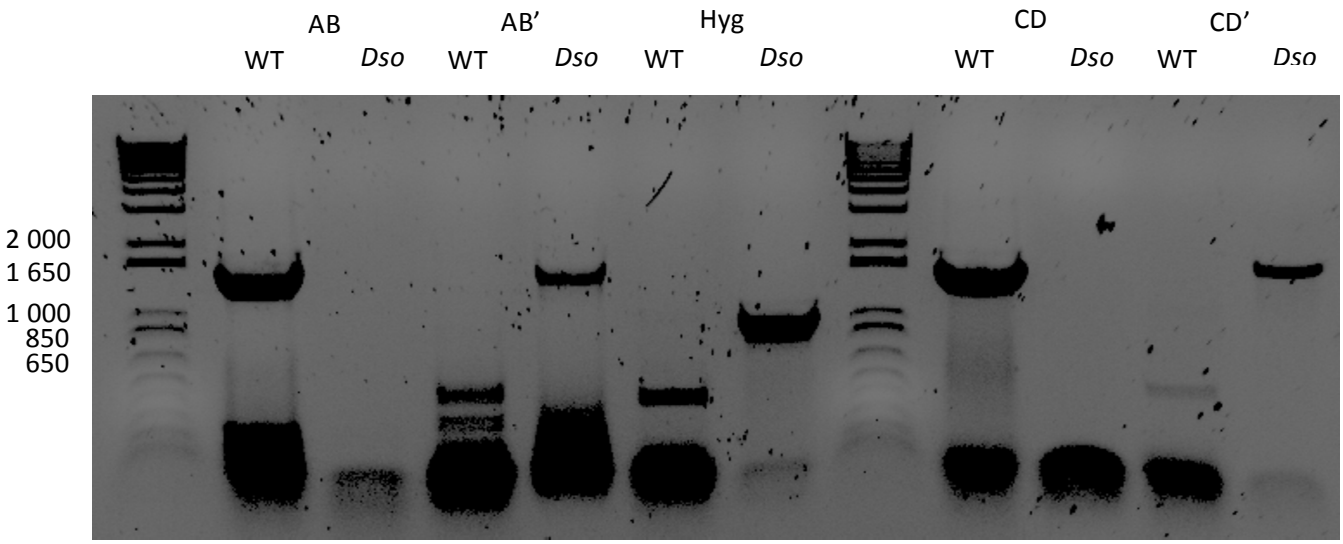

1-36 .MHIQLRND..R..Y.RNGY.grhcnPDN..YGTGSGYDDDSAAA
37-74 WLHIQLRNEHRR..Y.RSCW..whhdPDD..GRTGTGHDDDPAAA
75-111 RMQFRLRNECYR....NGYG..whhdPDD..HGTGSRHDDDSAAA
112-149 WLHIQLRDEHRR..Y.RSCW..whhdPDN..GRPGPRYKYDSAAA
150-186 WMYFRLRNE..R..Y.RNCY.eryrdPDD..YGTGSGYEYDGAAA
187-223 WMHIQLRDE..R..Y.RDCY.rrlydPDD..YGTGSGYEHDGTAA
224-261 WMHIQLRNKHYR..Y.RSCW..whydPDD..GRPGPGYKYDSAAA
262-298 RVYFRLCNECYR....NGYG..rhrnPDD..YGSGSRHDDDSAAT
299-335 WLYIQLRDE..R..Y.RDCY.rrlydPDD..YGTGSGHEHDGTAA
336-373 WMHIQLRNKHYR..Y.RSCW..whydPDD..GRPGPRYKYDSAAA
374-410 RVYFRLCNE..R..Y.RNCY.grlndPDD..YGTGSGYEHDSTAA
411-448 WMHIQLRNEYYR..Y.RSCW..whydPDD..GRPSPRYKYDSAAA
449-485 WVYFRLCNE..R..Y.RNCY.rrlydPDD..YGTGSGITTQTTTG
505-541 RVHCRLRDE..R..Y.RDGY.gwlydPDD..YGTGSGYEYDRTTA
542-579 WMHIQLHNEHYR..Y.RSCW..whydPDD..YGTGSGHEHDGTAA
580-619 WMHIQLRNEHYR..Y.RSCW..whydPDDddYGTGSGYEHNGTAA
620-656 WMYIQLRNEHYRngYgRHC......nPDD..YGTGSRHDDDCAAA
657-693 WMHIQLRDE..R..Y.RDCY.rrlydPDD..YGTGSGYKYNSATA
694-730 RVYIQLRNEHYRngYgRHC......nPDD..YGTGSRHDDDCAAA
731-767 WMHIQLRDE..R..Y.RDCY.rrlydPDD..YGTGSGYEHDGTAA
768-805 WMHIQLRNEHYR..Y.RSCW..whydPDD..YGTGSGHEYDGTAA
806-845 WMYIQLRNEHYR..Y.RSCWwhydsdDDD..YGTGSGYEHDGTAA
846-882 WMHIQLRNEHYRngYgRHC......nPDD..YGTGSRHDDDCAAA
883-914 WMHIQLRDE..R..Y.CDCY.rrlydPDD..YGTGSGYEH.....
916-931 ....DD..YGTGSGHEHDGTAA
932-969 WMHIQLPNEHYR..Y.RSCW..whydPDD..YGTGSGYEHDGTAA
970-1007 WMHIQLRNEHYR..Y.RSCW..whydPDD..YGTGSGYEHDGTAA
1008-1045 WMHIQLRNEHYH..N.RSCW..whydPDD..YGTGSGYKHDSTAA


A
200 µm
B
20 µm
C
D

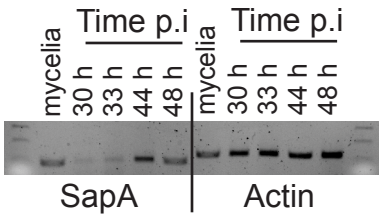


	Primer	Primer sequence
A	JEP2267	CGAGTCCAGCGACTTGCGAC
B	JEP2674	CGCACGCCGCGGCTGGCTTC
B'	JEP2266	CTTGTTGATGTACGGAGCACCC
C	JEP2675	CACATCTGGAGTACGGAAGG
C'	JEP2269	GCATGGACGAGCTGTACAAG
D	JEP2268	GTACCAGGTACCTACCCCTCC
HygR F	JEP2162	ATGCCTGAACTCACCGCGAC
HygR R	JEP2216	TTCCTTTGCCCTCGGACGAG

Primer pair	Expected size	Amplifies from WT	Amplifies from *Dso* knock-in
AB	1.6 kb	Yes	No
AB'	1.6 kb	No	Yes
CD	1.6 kb	Yes	No
CD'	1.6 kb	No	Yes
HygR F/R	1 kb	No	Yes



A



B

```
WBGene00004998|F08F1.6|SPP-13       MQSLQLLTFVLIALMVSVTFAEIIQKGKLAKHGMLKENKPNCLLTRSRLGCTCTTCKEIV
WBGene00004999|K09F5.3|SPP-14       MKT----TFCLLA-LVAVATAL--------------------ALPSQQDNAMSCLMCEVGV
WBGene00004991|T08A9.10|SPP-6       ------------------MTCM--------------------TGLR-QRNALDCEMCKLAV
WBGene00004987|T08A9.12|SPP-2       MKT----VLVLAA-VAAVALCV--------------------VLPK-ERSSLGCQMCELAV
WBGene00004990|T08A9.9|SPP-5        MKT----LLVLAV-LVAVASGL--------------------AIPE-NRSALSCQMCELVV
WBGene00005000|C48E7.10|SPP-15      -MNYLIL---LA--FLAITLTVSNV-------------EGAKKMHSDTSKPLCGLCVNVV
WBGene00004986|T07C4.4|SPP-1        -MTRILPCLFLV--LLAAAPLLANP-------------AN-PLNLKKHHGVFCDVCKALV
WBGene00004997|T22G5.7|SPP-12       MFSK--TVVVLL--MVVPAISLAQP-------------AS-PLVLKKSHGAFCHLCEDLI
                                                                           .     *  *   :

WBGene00004998|F08F1.6|SPP-13       NFTRMLILNHVPE-EQEVMEKVCYRIFGDDKKKESFCEELIKEELPDIIKYVRNHLEPKQ
WBGene00004999|K09F5.3|SPP-14       RAAENPADREAHT-VEDKFDAECKKELGIIPFAEKECEKYGNSKLDPIINELEGGTAPED
WBGene00004991|T08A9.10|SPP-6       KSADGDADKDTND-IKKDFDAKCKKAFHSIQFAPRECEHYVDKKLDPIIKELESGTSPKD
WBGene00004987|T08A9.12|SPP-2       KTYDGSADKDVTS-IKKDFDTECKKLFHAIPFAPQECEHYVNEKLDPIIKELESGTAPKD
WBGene00004990|T08A9.9|SPP-5        KKYEGSADKDANV-IKKDFDAECKKLFHTIPFGTRECDHYVNSKVDPIIHELEGGTAPKD
WBGene00005000|C48E7.10|SPP-15      KQLDQVLEHGGD--IEAAVDKFCKEDVP--SFMVDMCEKVIEKNLEYIINKLKDHEEADK
WBGene00004986|T07C4.4|SPP-1        EGGEKVGDDDLDAWLDVNIGTLCWTML---LPLHHECEEELKKVKKELKKDIENKDSPDK
WBGene00004997|T22G5.7|SPP-12       KDGKEAGDVALDVWLDEEIGSRCKDFG----VLASECFKELKVAEHDIWEAIDQEIPEDK
                                    .             .  .   *            * .  .     : . :      ..

WBGene00004998|F08F1.6|SPP-13       ACAKFC----------
WBGene00004999|K09F5.3|SPP-14       VCKKLKEC---------
WBGene00004991|T08A9.10|SPP-6       VCTKLGECS--------
WBGene00004987|T08A9.12|SPP-2       VCKKLGECPK-------
WBGene00004990|T08A9.9|SPP-5        VCTKLNECP--------
WBGene00005000|C48E7.10|SPP-15      ICTDILLCRTPKQYYFLETQK
WBGene00004986|T07C4.4|SPP-1        ACKDVDLC---------
WBGene00004997|T22G5.7|SPP-12       TCKEAKLC---------
```

Genome assembly, gene prediction and annotation.

As is standard practice, as far as possible, we adopted automated pipelines for data analysis. Judged against published fungal genomes, the genome of *D. coniospora* is clearly of a very high quality in terms of coverage and accuracy. We note, however, that these procedures did introduce a number of artefacts that will need correction in future versions of the annotated genome. Among the minor anomalies, we note the following:

As described in the main text, Scaffold43 was identified as being of mitochondrial origin. A more comprehensive investigation of this scaffold with TBLASTX against a database of mitochondrial genes identified three mitochondrial-specific ATP synthases (*atp6*, *atp8*, *atp9*), one cytochrome b (*cob*), two cytochrome oxidases (*cox1* and *cox2*), six subunits of NADH dehydrogenase (*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*), and one *rps5* ribosomal protein. Fungal genomes generally have a conserved structure, with 14 protein-encoding genes arranged in the order *nad2*, *nad3*, *atp9*, *cox2*, *nad4L*, *nad5*, *cob*, *cox1*, *nad1*, *nad4*, *atp8*, *atp6*, *cox3* and *nad6*. The absence of genes for *cox3* and *nad6* from Scaffold43 suggested a truncation of the sequence. Genes for *cox3* and *nad6* were found next to each other (coordinates 2196045-2196848 and 2197076-2197738) within the sequence of omap49267b (a scaffold of 10013798 bp), probably the result of an assembly error.
We also observed that a short DNA sequence was duplicated on 2 separate unmapped scaffolds. Specifically, scaf62 (5655 bp) was found to be 99% identical to scaf63 (5653 bp). Further, the 5' 3334 bp of scaf62 matches perfectly the start of both scaf59 and scaf60, while the 3' 1939 bp matches a sequence within omap3956. The most plausible explanation for all these observations is an error in automated assembly. To avoid biasing slightly our analyses, the 8 supernumerary copies of the proteins predicted for these regions were removed from an original Augustus-derived set of 8742 proteins, together with a single protein that had been predicted despite its lack of a 5' ATG. There may be other isolated examples of misassembly leading to artificial duplication of genes. For example, we note that the 3 "glycine rich protein" (PF12810) domain genes referred to in the main text (fg9098.t1, fg9056.t1, fg8010.t1) on omap6109, scaf41 and scaf68 are 100% identical at the DNA level, and so could be the result of very recent gene duplication events but might be the consequence of imperfect assembly. At the level of individual proteins, a detailed analysis of *D. coniospora* CAZy proteins (Supplementary Table S8) suggested that 16/240 (<7%) of predicted gene models might not be entirely correct. Rectification of these different types of errors will require painstaking re-analysis and expert curation and will be included in a future version of the genome.

With regards 3rd party annotation, among the domains unique to *D. coniospora*, PF02609 (Exonuclease VII small subunit) was present in g4632.t1. This is a clear homolog of the highly conserved protein Rad50 (Supplementary Table S5), a protein devoid of a PF02609 domain. Further inspection showed the domain assignment for g4632.t1 to be very tenuous, and due to the insertion of a hexapeptide sequence "ERYRQD" within the otherwise well-conserved 200 residue-long AAA (PF13476) domain. The PF02609 domain is therefore not discussed further in the main text.

Note added in proof.

While the revised version of this article was in review, a genome sequence for *D. coniospora* strain ARSEF 6962 was reported [1]. As explained below, these strains were derived from a common isolate. Despite global concurrence between the 2 sequences and between the sets of predicted genes, preliminary comparisons indicate an unexpectedly high level of sequence divergence, including an elevated level of single nucleotide polymorphisms, as well as multiple indels, including some of up to several kb in length. Determining whether these reflect errors in sequencing and/or assembly for one or other sequence, or true molecular divergence will require targeted investigation in the future.

One striking difference concerns the number of predicted chromosomes, 9 (this study) versus 3 [1]. As optical mapping was used in both cases, a detailed comparison of the respective data will be required to provide an explanation for this discrepancy.

Another notable difference is the prediction from the ARSEF 6962 sequence of a protein containing a MATα (PF04769) domain. This lead Zhang *et al*. to speculate on the existence of a cryptic sexual cycle for *D. coniospora* and to suggest that it might be a heterothallic species [1]. Again, this important point will require clarification.

The strain used in the current study was originally provided to us in 1998 by Hans-Börje Jansson, then at Lund University, Sweden. At the time, he wrote that it was the strain "called "ny" ATCC# 96282 used in the JON paper [Journal of Nematology (26:430; 1994)]". In the "JON paper", one finds mention of "The fungus *Drechmeria coniospora*, designated isolate no. 5 by Jansson (9)", with reference 9 being:

Jansson HB. Adhesion to Nematodes of Conidia from the Nematophagous Fungus *Drechmeria coniospora*. Journal of General Microbiology. 1993;139:1899-906.

The strain was deposited in the CBS-KNAW culture collection by Jansson and is held under the name CBS 615.82, and is in the ARSEF collection as strain 2468. To summarise, ATCC 96282 is variously called, "ny", "isolate no. 5", "CBS 615.82" and "ARSEF 2468". The strain ARSEF 6962 is described in the ARSEF catalogue as a "reisolation of ARSEF 2468", and should therefore be a re-isolation of ATCC 96282. It is important to note that the strain provided to us by Jansson had been in almost continuous culture in our laboratory, passaged through infection of *C. elegans*, for some 15 years before a sample of its DNA was extracted for genome sequencing.

1. Zhang L, Zhou Z, Guo Q, Fokkens L, Miskei M, Pocsi I, et al. Insights into Adaptations to a Near-Obligate Nematode Endoparasitic Lifestyle from the Finished Genome of *Drechmeria coniospora*. Sci Rep. 2016;6:23122. doi: 10.1038/srep23122. PubMed PMID: 26975455; PubMed Central PMCID: PMCPMC4792172.

Supplementary Methods

Media

Nematode Growth Medium (NGM) [1]:

NaCl	3 g
BactoPeptone	2.5 g
BactoAgar	20 g
5 mg/mL cholesterol in EtOH	1 mL
1 M $MgSO_4$	1 mL
1 M $CaCl_2$	1 mL
1 M KPO_4 pH 6.0	25 mL

Nematode Growth Medium with Yeast extract (NGMY). For 1 litre:

NaCl	3 g
BactoPeptone	2.5 g
Yeast Extract	20 g
5 mg/mL cholesterol in EtOH	1 mL
1 M $MgSO_4$	1 mL
1 M $CaCl_2$	1 mL
1 M KPO_4 pH 6.0	25 mL
100 mg/mL Ampicilin	1 mL
Gentamycin	1 mL

Solid NGMY medium was made following the same recipe but with the addition of 20 g/l BactoAgar.

Plasmid construction

To construct the pLH4237 plasmid containing *hphgfp* under the control of the *D. coniospora* beta-tubulin promoter (from g807.t1), the backbone, *hphgfp* and *trpC* terminator sequences from pPK2*hphgfp* [2] were amplified with primers JEP2162/JEP 2163 and fused with *D. coniospora* beta-tubulin promoter (PCR amplified from genomic DNA with JEP2160/JEP2161) via Gibson assembly [3].

To construct the plasmid targeting *Dso* (g1469.t1), the gene together with 1 kb 5' and 3' flanking arms were PCR amplified from genomic DNA (using primers JEP2239/JEP2240) and cloned into pGEM-T Easy (Promega). This plasmid was then used as a template to amplify the 5' and 3' flanking arms and the pGEM-T backbone sequence (using primers JEP2237/JEP2238) and ligated to the beta-tubulin promoter *hphgfp* sequence from pLH4237 (PCR amplified using primers JEP2242/JEP 2243) by Gibson assembly to give pLH4256.

To make the SapA::dsRed reporter, we PCR amplified 4 separate DNA fragments: (i) the *sapA* coding region without the stop codon and with 1 kb of sequence upstream of the predicted ATG, from genomic DNA with primers JEP2594/JEP2596. (ii) The coding sequence for dsRed from the plasmid pLH4238 with JEP2592/JEP2593. (iii) The beta-tubulin promoter, *hphgfp* and *trpC* terminator from pLH4237, with JEP2242/JEP2591 (iv) The 5' and 3' flanking arms of *Dso* and the pGEM-T backbone sequence with JEP2237/JEP2238. The fragments were joined by Gibson assembly to give pLH4244.

Primer sequences

JEP2160 GATATCGAGCTCGGTACCCGGCTAGGGTGCTCCGTACATC

JEP2161

GTCGCGGTGAGTTCAGGCATCTATATCGAAGAAGGAAACTCA

AG

JEP2162 ATGCCTGAACTCACCGCGAC

JEP2163	CCGGGTACCGAGCTCGATATC
JEP2237	TGAAGGATGAGCGACGGCGC
JEP2238	CATCGTCCCTTGCTCGCACAG
JEP2239	AGCTGTGCTCCGTGTGTTGT
JEP2240	CTCCGCCAGAAATGCACCGT
JEP2242	CTGTGCGAGCAAGGGACGATGTAGGGTGCTCCGTACATC
JEP2243	GCGCCGTCGCTCATCCTTCACTTGTACAGCTCGTCCATGC
JEP2592	GGGCCCATGGCCTCCTCCGA
JEP2593	GCGCCGTCGCTCATCCTTCATCAGTTGGAATTCG
JEP2594	CTCCACTCGACCTGCAGGTCGAATGGCCCTCCAAGTTG
JEP2596	GGAGGAGGCCATGGGCCCTGTCGTGGTAGCAGCCGACG
JEP2242	CTGTGCGAGCAAGGGACGATGTAGGGTGCTCCGTACATC
JEP2591	CTGCAGGTCGAGTGGAGATG

1. Stiernagle T. Maintenance of *C. elegans*. http://www.wormbook.org: The *C. elegans* Research Community ed; 2006. Available from: http://www.wormbook.org.
2. Michielse CB, van Wijk R, Reijnen L, Cornelissen BJ, Rep M. Insight into the molecular requirements for pathogenicity of *Fusarium oxysporum* f. sp. *lycopersici* through large-scale insertional mutagenesis. Genome biology. 2009;10(1):R4. Epub 2009/01/13. doi: 10.1186/gb-2009-10-1-r4. PubMed PMID: 19134172; PubMed Central PMCID: PMC2687792.
3. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature methods. 2009;6(5):343-5. Epub 2009/04/14. doi: 10.1038/nmeth.1318. PubMed PMID: 19363495.

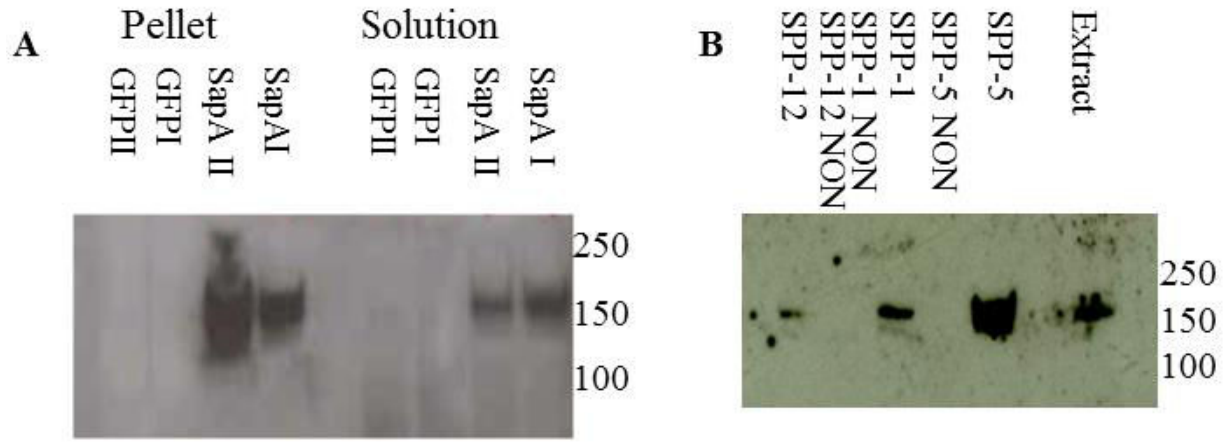


Figure 4.1 Immunoprecipitation of SapA-SapB complex. (A). Comparison of fungal protein extraction protocols. Insoluble (Pellet) and soluble (Solution) fractions of proteins extracted from fungal strains expressing SapA-DsRed (SapA) or HPH::GFP (GFP) using one of 2 protocols (I and II). Proteins were separated by SDS-PAGE and after transfer to a membrane revealed with an anti-RFP antibody. (B). Extracted fungal proteins were mixed with the different purified recombinant His-tagged SPP proteins SPP-1, SPP-5, SPP-12; anti-DsRed or non-specific binding (without the anti-DsRed antibody) beads were used to pull down SapA-DsRed from the mixture; Western blotting was stained with anti-DsRed antibody; The crude fungal extract is indicated by "Extract", the protein extracted by anti-DsRed beads defined with SPP-1, SPP-5 or SPP-12; the protein retained on non-specific beads indicated by the suffix "NON".

4.2 Additional information I- SapA-SapB protein interaction

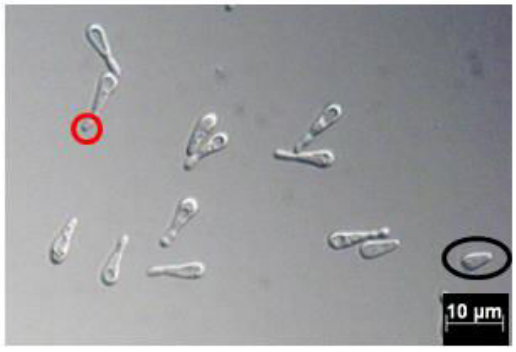
One of the most interesting potential virulence factors we found from the bioinformatics analysis was SapA (g3895), characterised by its 3 SapA domains (PF02199). Not only were no orthologous found in the 11 other fungi analysed, but only 3 fungal proteins containing PF02199 are reported in PFAM, all from *Fusarium* spp. Its structure immediately suggested a possible regulatory interaction with SapB-domain protein(s). In *C. elegans*, these form part of the large SPP family, also called caenopores (ROEDER *et al.* 2010). We assayed the interaction between SapA and the nematode caenopores both *in vitro* and *in vivo*. To do so, we needed to be able to extract SapA from the surface of spores in strain (JEM907) where it is highly expressed (see 4.1). We therefore tested two protocols for their ability to extract fungal protein from spores. With both methods we could extract the fungal protein; we chose method I, as it generated comparatively more soluble protein (figure 4.1 A).

We launched an *in vitro* assay to determine the affinity of SapA with three purified recombinant worm SapB proteins SPP-1, SPP-5 and SPP-12. Firstly, we incubated total protein extracts from JEM907 with these three His-tagged SPP proteins. Then, pulled down the protein complexed to DsRed-SapA using anti-RFP beads. These protein complexes were run on SDS-PAGE, then transferred to nitrocellulose membrane and stained with anti-His antibody to check for the presence of the different SPP proteins. As a preliminary step, we had confirmed the specificity of anti-RFP for pulling down the DsRed containing protein complexes, as a DsRed-SapA band was present in all samples immunoprecipitated with anti-RFP beads but not when non-specific beads were used (figure 4.1 B). We subsequently proved that G3895 can bind SPP-5 *in vitro* (see section 4.1).

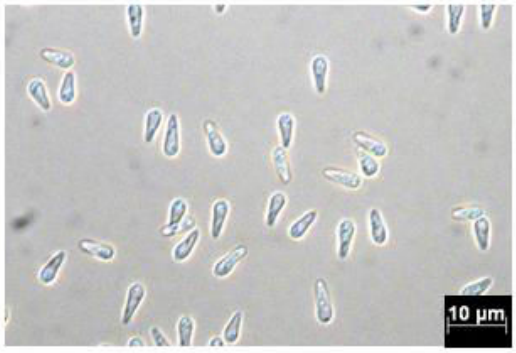
Table 4.1 Worm proteins found in JEM907-infected worms

Gene	Function
D1054.10	Unknown function, worm specific
C34F6.8	Isocitrate dehydrogenase
T01C3.7	rRNA 2'-O-methyltransferase fibrillarin
T02G5.9	Lysine-tRNA ligase

A



ATCC 96282, JE lab strain



ARSEF 6962, Zhang's strain

B

	ATCC 96282	ARSEF 6962
Length (um, average), Standard deviation	9.1, 2.01	5.6, 0.67
With bud% (more than 80 spore counted)	71.4	<1

Figure 4.2 Spore morphology difference between ATCC 96282 and ARSEF 6962. (A). Spores after 1 month of NGM culture; image for ATCC 96282 was taken with Nomarski microscopy, image for ARSEF 6962 was taken with bright filed microscopy; the red circle indicates an adhesive bud, the black circle a spore without a bud. (B). Two strains differs on their spores: average length and percentage of spores with bud.

In our *in vitro* experiments, we were only able to test the SPP proteins that M. Lieppe (Univ. Kiel) kindly provided. We wished to determine whether SPP-5 was equally a binding partner of sapA *in vivo*. We therefore took a non-biased approach based on MS to determine the host proteins that interact with fungal SapA. We used worms infected with JEM907. The SapA-DsRed and proteins interacting with it were pulled down with anti-RFP beads after 49 h of infection, and sent for MS analysis. In a first round of MS analysis, out of a total of 1215 identified proteins, only 4 worm proteins were only found in the sample from JEM907 infected worms; none corresponded to a SPP protein (Table 4.1), indeed no SPP proteins were identified at all. The failure to identify worm proteins that interact with sapA *in vivo* may reflect a technical issue with the MS; we got a comparatively limited number of peptides from the analysis. We are planning to repeat the experiment in the future.

4.3 Additional information II- Comparison between ATCC 96282 and ARSEF 6962

A second paper reporting a genome sequence for *D. coniospora* was published almost at the same time as ours. It is based on ARSEF 6962 (ZHANG *et al.* 2016), a strain derived from CBS 615.82, isolated from Denmark in 1982 at the latest, and distinct from the one ATCC 96282 we are using that was isolated in Sweden in 1990 (see http://journals.plos.org/plosgenetics/article/comments?id=10.1371/journal.pgen.1006017). We compared the morphology of these two strains. We did not see any marked difference with regards to their colony shape and hyphal growth, but their spores on NGM plates can be distinguished (Figure 4.2A). Our strain had an average length of 9.1 µm and ARSEF 6962 was around 5.6 µm. And more importantly, under the conditions of the assay, more than 70% of our spores had adhesive buds, in contrast with less than 1% of ARSEF 6962 (Figure 4.2 B). The bud has been reported to be a transient structure (JANSSON AND FRIMAN 1999) essential for spore

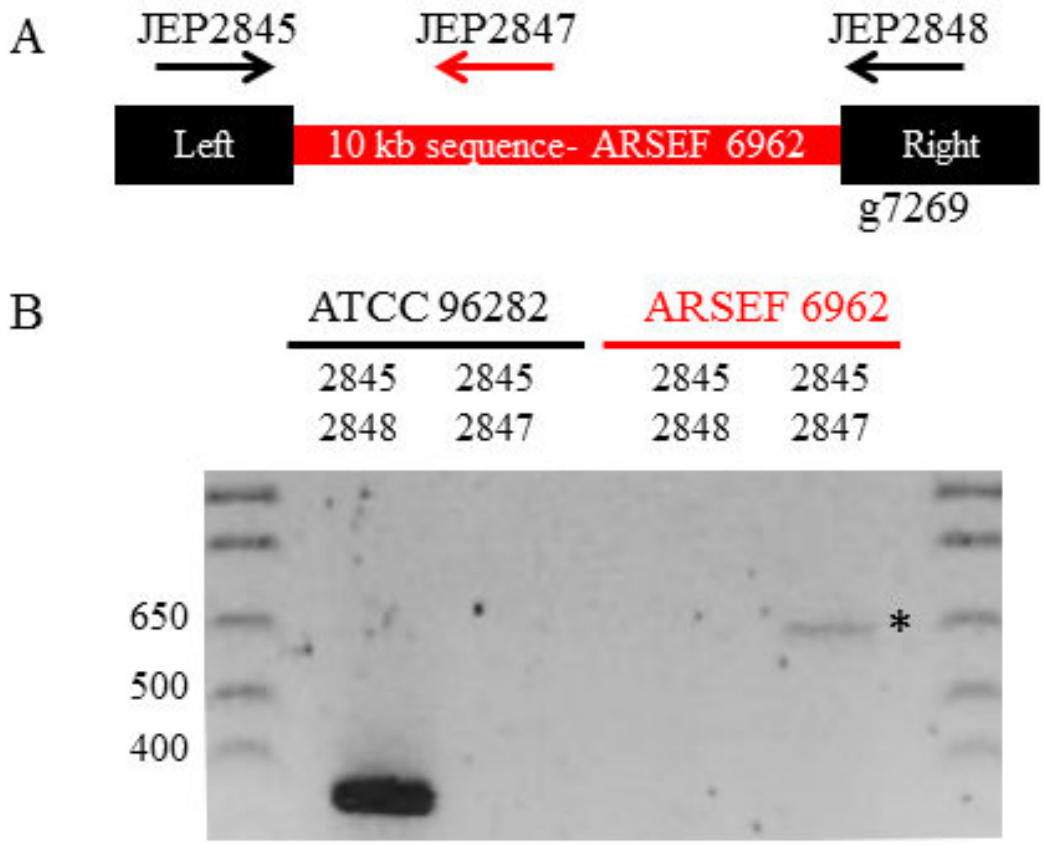


Figure 4.3 PCR verification for the extra 10 kb sequence in ARSEF 6962 genome. (A). the location of the extra 10 kb sequence of ARSEF6962 in ATCC 96282; JEP primers indicated were used for subsequent PCR confirmation. (B). Using DNA from ATCC 96282, we obtained the expected small band using the primers JEP2545/JEP2548 but did not see the extra sequence-specific band (with JEP2545/JEP2547). For ARSEF 6962, the opposite was true. The PCR products were sequenced and they are in accordance with the expected genomic sequences.

binding to the worm cuticle. The observed difference may indicate that ARSEF 6962 adhesive buds mature at a different rate from ATCC 96282, or less likely, that the former uses a different strategy to infect its host. It should be noted that ATCC 96282 has been maintained in continuous culture on NGM plates for more than 10 years, so may have undergone changes to its growth patterns. This merits further investigation.

Intriguingly, ARSEF 6962 was reported to have just 3 chromosomes, whereas we reported 9. Both analyses involved optical mapping that is supposed to give a direct visual measure of chromosome number. It is, however, an imprecise method. Several chromosomes that are reported to be separate in ATCC 96282 but concatenated in the ARSEF 6962 may have been joined together in the latter sequence by the inclusion of very long stretches of (0.5Mb) "N" (unknown) nucleotides. On the other hand, this latter sequence used PacBio reads that are substantially longer than those used for the assembly of the ATCC 96282 sequence. A detailed comparison of the two genomes is on-going. Already, we have defined several differences on the genomic level through a whole genome alignment. For example, ARSEF 6962 contains an extra 10 kb of sequence when compared to ATCC 96282 genome's 5' of g7269 (Figure 4.3 A). A PCR assay was performed and confirmed the existence of this extra sequence in ARSEF 6962 (Figure 4.3 B). This particular region does not contain any predicted proteins. It will be interesting in the future to identify indels and other polymorphisms that potentially affect genes involved in virulence.

4.4 Using APEX to identify fungal effectors

The APEX method (REINKE *et al.* 2016) relies on the delivery of biotin-phenol to the host cells. As mentioned in the introduction, we used *bus-8* to facilitate this. *bus-8(RNAi)* should provoke a skiddy phenotype (GRAVATO-NOBRE *et al.* 2005; PARTRIDGE *et al.* 2008); we observed this

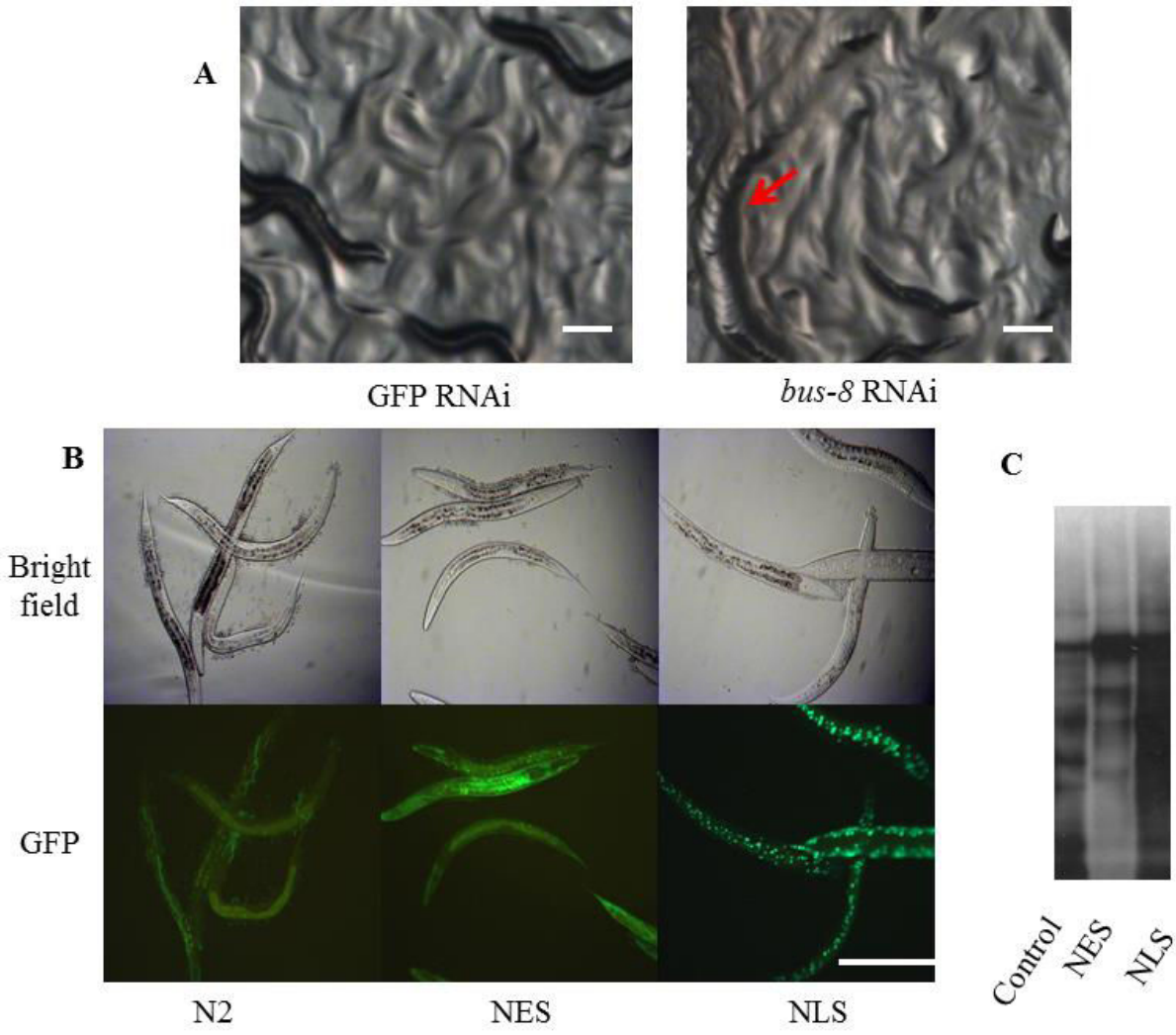


Figure 4.4 Feeding and washing method for sample preparation. (A). Synchronized L1 stage N2 worm fed with GFP and *bus-8* RNAi bacteria for 48 h; *bus-8(RNAi)* fed worms exhibited a skiddy phenotype (red arrow), bar = 200 µm. (B). Infected worms after M9T /9 (formula see REINKE *et al.* 2016) wash; strains were N2, NES (*dpy-7p*::APEX::GFP::NES), NLS (*dpy-7p*::APEX:GFP::NLS), bar = 200 µm. (C). Silver staining for the total protein immunoprecipitation; control strain was *spp-5* promoter::GFP, NES and NLS were the same with B; worms were infected for 44 h before extraction and proteins eluted from streptavidin beads were loaded on the gel.

under our experimental conditions (figure 4.4 A). The initial protocol provided by Emily Troemel's lab used harsh conditions for the washing step that disrupted *D. coniospora* infected worms, before the biotin-phenol biochemical reaction. In order to obtain intact worms, we modified the protocol using a more gentle method (refer to 5.4). We showed that the worms were still intact after the wash step as the APEX::GFP chimera is well located inside the epidermal cytosol or nuclei (figure 4.4 B).

Next, the biotinylated proteins from samples of infected worms were extracted, and purified with streptavidin beads, loaded on the SDS gel and silver stained. We showed that the biotinylated protein in infected worms was enriched in the NLS APEX strain but not in the NES strain compared to the control strain (Figure 4.4 C). The signal in the control strain is higher than NES sample reflects endogenous biotinylation. The experiment was repeated for 3 times and similar results were obtained.

The three batches of samples were analysed with MS. To verify the efficiency of APEX strains used in this protocol, we performed a protein localization analysis on the total protein information we got from MS. As APEX were supposed to label the proteins locally in the worm epidermal cytosol in the NES strain or the nuclei in the NLS strain, we should be able to see more epidermal proteins in both NES and NLS samples, and more nuclear protein in the NLS sample. Analysis of predicted protein localization on the 2nd round MS showed no enrichment of epidermal proteins in NES or NLS samples: both had around 12%, not markedly different from N2. Similarly, nuclear proteins were not enriched in the NLS sample (Figure 4.5). We got similar result for the first round MS result analysis. The high number of total N2 protein reads indicates a strong biotinylation background in *C. elegans*. The similarity on predicted protein localization

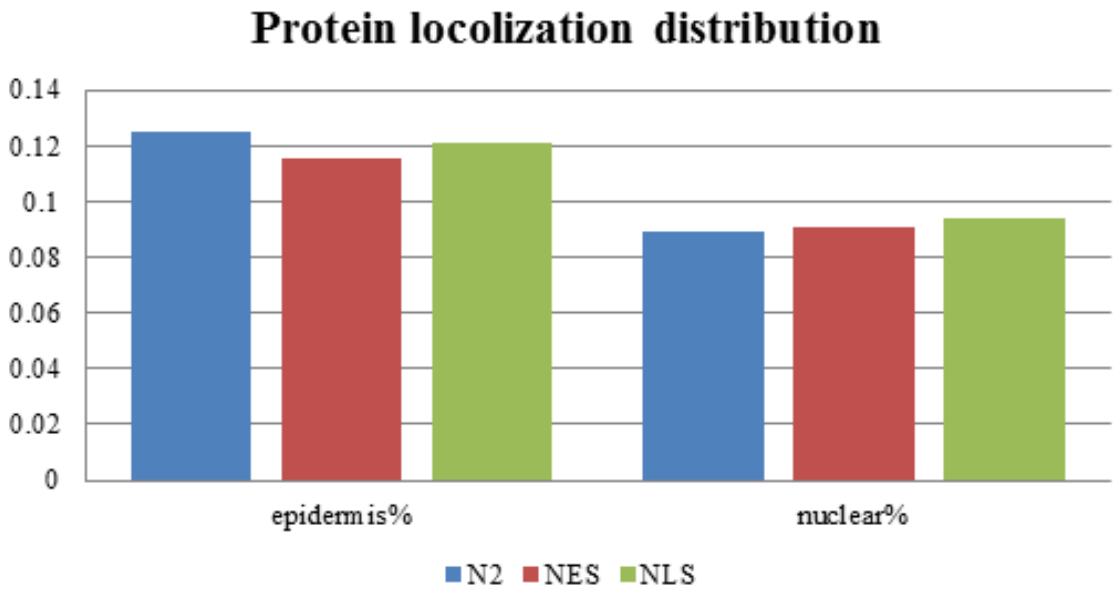


Figure 4.5 Protein localization from the 2nd round of mass spectrometry analysis. NES, NLS represent the strains indicated in Figure 4.4. The nuclear and epidermis protein were pulled out from all sample and divided by the total proteins, the resulting proportion was used to make the columns.

Table 4.2 Three rounds of Mass spectrometry analysis

Round	Total fungal protein*	Specific fungal protein*
1st	17	6
2nd	15	8
3rd	12	none

*For each round of MS, samples from NLS, NES and non APEX strains, infected or not infected, were analyzed. Total fungal protein included fungal proteins from NLS, NES and non APEX sample. Total proteins were fungal proteins from NES, NLS and non APEX samples; specific fungal protein is proteins only found in NES or NLS sample.

between NES, NLS and N2 suggests that in our hands the APEX method is not efficient for labelling the protein in a specific cell compartment. Despite the apparent limitations of the method, the MS analysis did reveal the presence of a small number of fungal proteins; a total of 14 specific fungal proteins were pulled out from NLS or NES samples (Table 4.2). No specific fungal protein was obtained from the 3rd round due to a technical issue.

We further analysed the candidate proteins using the secretome data and gene expression information from *in vitro* cultured mycelium and spore samples (data refer to 4.1). An ideal candidate would be one that was expressed at a low level in the *in vitro* cultured fungus and secreted. We found, however, that 11 out of 14 proteins were highly expressed in the *in vitro* cultured fungus compared with the tubulin gene g3737; further, none of them was predicted to be secreted from the fungus. This left 3 genes which are still potentially interesting because of their low expression level, g5131 (Mitochondrial PGP phosphatase domain), g3166 (Heat shock factor binding domain) and fg8975 (*D. coniospora* specific protein, absent from ARSEF 6962). They may merit further study.

Table 4.3 Proteins from Mass spectrometry analysis

Gene	Pfam	Expression (read counts)		Secreted	Round
		Mycelium	Spore		
g3737	PF00091|Tubulin/FtsZ family, GTPase domain(1.9E-69)	11057	20066	No	-
g5391	PF00043|Glutathione S-transferase, C-terminal domain(9.7E-15)	29714	39714	No	1
g5175	PF00071|Ras family(5.0E-56)	2943	3220	No	1
g712	PF01248|Ribosomal protein L7Ae/L30e/S12e/Gadd45 family(3.9E-21)	21584	32074	No	1
g5131	PF09419|Mitochondrial PGP phosphatase(3.9E-64)	87	99	No	1
g5691	PF00009|Elongation factor Tu GTP binding domain(1.3E-45)	150474	200327	No	1
fg8975	No Pfam domain, *D. coniospora* specific	144	155	No	1
g7853	PF00012|Hsp70 protein(7.5E-229)	33583	55313	No	2
g1022	PF00702|haloacid dehalogenase-like hydrolase(2.6E-22);	111960	70367	No	2
g3745	PF01201|Ribosomal protein S8e(2.0E-46)	18196	29744	No	2
g3166	PF06825|Heat shock factor binding protein 1(7.1E-6)	842	659	No	2
g4282	PF00183|Hsp90 protein(7.4E-255)	200469	220242	No	2
g1429	PF00270|DEAD/DEAH box helicase(2.6E-44)	33019	58144	No	2
g709	PF01293|Phosphoenolpyruvate carboxykinase(3.7E-214)	38820	27985	No	2
g178	PF00297|Ribosomal protein L3(2.9E-117)	36892	59209	No	2

4.4.1 Discussion and perspective

Our study with strain ATCC 96282 together with Zhang's on strain ARSEF 6962 (ZHANG *et al.* 2016) gave a molecular insight into *D. coniospora*. Both studies concluded that *D. coniospora* belongs to Ophiocordycipitaceae, currently most closely related, among sequenced species, to the truffle pathogen *Tolypocladium inflatum*. *D. coniospora* has adopted an obligatory pathogenic lifestyle, and has lost many genes needed for a saprophytic lifestyle. There are some obvious differences between these two strains' genomes, the most dramatic being the predicted number of chromosomes, 3 in ARSEF 6962 versus 9 in our sequence data. It is inconceivable that the two strains should really be so different; we are currently attempting to resolve the issue. Nevertheless, the strains are distinct. At the molecular level, for example, I validated one representative sequence difference at the 5' side of g7269 in ATCC 96282; at the equivalent genomic position in ARSEF 6962, there is a ca. 10 kb insertion of extra sequence not found in ATCC 96282. An in-depth genome-wide gene comparison is currently underway. I also observed morphological differences between their spores. Given the divergence of these two strains on their genome structure and morphology, it will be interesting to see whether there are differences in the way in which they infect their host, at the cellular and molecular levels.

Comparative genomic analysis with other nematophagous fungi, such as *A. oligospora*, *P. chlamydosporia* and *H. minnesotensis*, revealed the unique molecular basis for the pathogenic lifestyle of *D. coniospora*. The species-specific genes reflect its interaction with its hosts and determine the fungal lifestyle. Zhang *et al.* concluded that half of the *D. coniospora* small secreted proteins (124 of 257) are species-specific (ZHANG *et al.* 2016). Similarly, we defined 16.3% of all proteins as being *D. coniospora*-specific, on the basis of currently available sequences. These species-specific proteins were considered as potential effectors that might

contribute to *D. coniospora* pathogenesis. We experimentally addressed this assumption by studying the function of one *D. coniospora*-specific SapA domain containing protein (G3895). We demonstrated that G3895 is secreted from the fungus and it can probably inhibit host immune activity by interacting with *C. elegans* caenopores, with SPP-5 currently being a candidate. G3895 was highly expressed in the later stage of infection, which suggests that this protein may have other functions. It is interesting to note that the SapA domain was reported to be essential for prosaposin and pulmonary surfactant-associated protein B transport (LIN *et al.* 1996; LEFRANCOIS *et al.* 2002). In this context, we need to identify the host proteins that interact with G3895. Similarly, other potential virulence effectors also need further functional verification.

To identify directly potential fungal effectors, we used an APEX based method to fish out pathogenic proteins that were secreted into the host. We were disappointed by the results. Very recently, with this method Reinke *et al.* reported the identification of 82 microsporidial proteins from the intestinal cells of infected worms. These proteins were quickly evolving, grouped into large *Nematocida*-specific gene families, and enriched in targeting signals. Reinke *et al.* also showed that this method preferentially labelled pathogen proteins which were secreted into the host. Most of the genes corresponding to the proteins they identified were moderately expressed during infection (REINKE *et al.* 2016). In our case, 14 proteins were identified from the infected worms, 3 of which are lowly expressed in *in vitro* cultured fungus. But, we did not see any enrichment for nuclear proteins in the NLS samples compared to NES samples. The limited number of pathogen proteins identified from our sample may be because our MS coverage is 20 times lower than Reinke's, a total of 2,171 peptides per sample against 43,499 peptides. This is in accordance with the comparison of peptides counts for the 74 proteins present in both data sets.

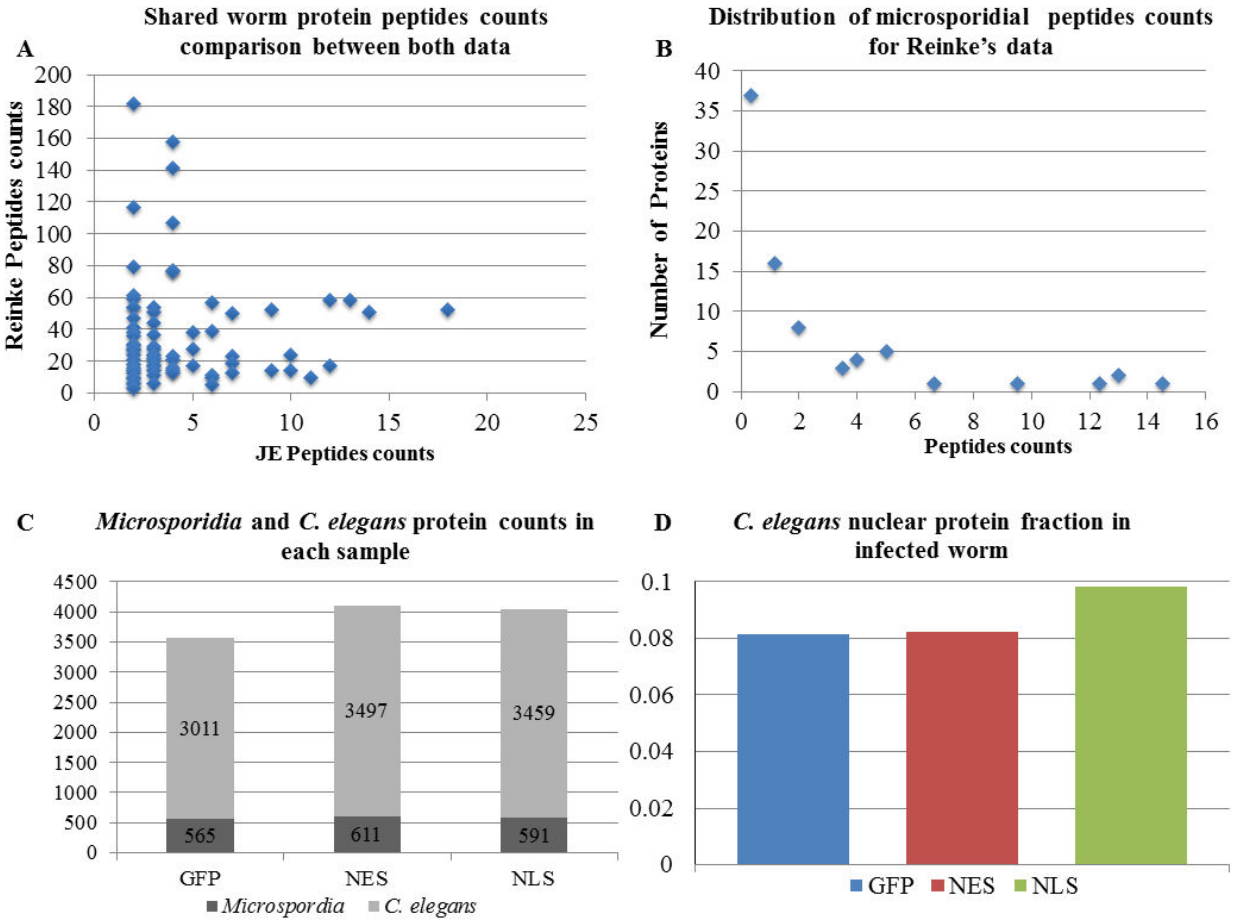


Figure 4.6 Peptide counts from a published APEX study. The data in Reinke's MS shows higher peptide counts and low pathogenic peptide coverage. (A). Peptide counts comparison of proteins identified in our data and Reinke's. We compared the average peptide counts of 72 proteins found in both datasets (for list, see Material and Methods); all proteins are represented by less than 20 peptides on JE side, 66% of proteins by more than 20 peptides from the Reinke study. (B). Distribution of peptide counts for microsporidial proteins from Reinke's data. The average counts for 79 microsporidial proteins found only in NES or NLS samples in Reinke's data were plotted. (C). The number of *Microsporidia* and *C. elegans* proteins identified in the 3 samples. The proteins from the GFP sample represent the non-specific signal. (D). The fraction of *C. elegans* nuclear proteins. The number of predicted worm nuclear proteins was divided by the total number of worm proteins identified in each of the 3 samples.

In our data none of these proteins had more than 20 peptide counts while the majority (64%) were represented by more than 20 peptides in Reinke's data (Figure 4.6 A). In addition, despite the depth of the MS coverage, the host-exposed microsporidial proteins were still poorly covered. Within the infected NES or NLS samples, more than 90% of these microsporidia proteins had an average peptide count less than that of the entire host and pathogen proteomes (10.2 peptides; Figure 4.6 B). There was also an extremely high background signal. The GFP control sample gave almost as many proteins (both pathogen and host-derived) as the NES or NLS experimental samples (Figure 4.6 C). Indeed, the percentage of microsporidial proteins was slightly decreased in the NLS, NES samples compared to the control (NES 14.9%, NLS 14.6% versus GFP 15.8%). We also calculated the percentage of *C. elegans* nuclear protein in each sample. Although the nuclear protein fraction was increased in the NLS sample, however, the difference was marginal (Figure 4.6 D). The underrepresentation of nuclear protein in our NLS sample further suggested the strong contribution of endogenous biotinylation, as has also been reported in other similar systems (OOI *et al.* 2010; SCHAFFER *et al.* 2010). In conclusion, because of the strong biotinylation the current APEX based method is not efficient for fishing the fungal virulence effectors unless we increase greatly the depth of the MS analysis.

Apart from the biochemical method, transcriptomic analysis of samples from infected worms can also give insights into fungal pathogenesis. In our case, fungal transcripts represented less than 0.002% of all transcripts sequenced from samples of worms infected for 5 and 12 hour (from 77 million and 123 million RNAseq reads), corresponding to just 537 fungal genes. Nevertheless, these genes provide an insight into not only the physiological state of the fungus but also early fungal infection strategies. For example, the chitinase protein G6474 is probably used by the fungus to digest the worm chitin-rich cuticle during the beginning stage of infection.

The very low coverage of the fungal transcriptome may be due to the fact that relatively few hypha are inside the worm at the two time points. Further, the fungal cell wall is relatively hard to destroy just with Trizol®. To overcome these obstacles, we can try to use a harsher way to extract the total RNA, such as using Fastprep® which is normally used to extract filamentous fungal DNA or RNA. In Zhang's paper, they took samples from worms with 8 day of infection for RNA sequencing (ZHANG *et al.* 2016). Considering that in normal conditions worms die after 48-60 h of infection (COUILLAULT *et al.* 2004), the relevance of such a late time point is debatable. It is possible, however, that their fungus which is cultured on a liver-kidney medium partial lost its virulence; spores from worm-free media were found to be less infective (LOHMAN AND SIKORA 1989). Unfortunately, no survival curves for their isolate were presented. We should conduct direct comparisons of the virulence of the two strains in the future.

D. coniospora could be defined as a hemibiotrophic fungal pathogen since its lifestyle is different at the early and late stages of infection. From the stand point of worm immunity, the necrosis-associated genes *asp-3*, *asp-4* can be triggered by *E. faecalis*, *E. carotovora* and *P. luminescens* infection, which we did not see at the beginning stage of *D. coniospora* infection (WONG *et al.* 2007; ENGELMANN *et al.* 2011). This shows that *D. coniospora* resembles a hemibiotroph infect the host without triggering its cell necrosis at this stage. Previous electronic microscopic studies provide a detailed morphological picture of the consequence of the devastating *D. coniospora* growth inside of worms at late stage of infection (DIJKSTERHUIS *et al.* 1991). In parallel with understanding how the fungus triggers and then potentially weakens the worm's immunity at the early stage of infection, the later stages merit further investigation in the future. It will be very interesting to see how the fungus switches from biotrophic to necrotic lifestyle in the late stage of infection. On the other hand, at the later stage, the fungal necrotic

effectors are going to change the host immune reaction drastically; it will be worthwhile to identify these effectors and to study how these effectors are functioning inside the host.

This chapter describes materials and methods not detailed in the published articles.

CHAPTER 5. MATERIALS AND METHODS

5.1 Plasmid construction based on Gibson assembly and *ccdB* selection

Media, enzymes and chemicals

For the antibiotic selection, ampicillin (100ug/ml) was added to LB plates or liquid. Gibson assembly® Master Mix, restriction enzymes *NotI*, *ClaI*, *NarI* were purchased from New England Biolabs. For amplifying the fungal genomic sequence, Takara PrimeStar HS DNA Polymerase was used. Promega GoTaq® Master Mix was used to verify the resulting plasmid. PCR product was purified with QIAquick PCR Purification Kit. Plasmid was purified with HiSpeed Plasmid Kits.

Bacteria strains

DB3.1 *E. coli* strain was used to amplify *ccdB* containing plasmid. NEB 5-alfa was used to amplify and maintain the other plasmid, and was purchased from New England Biolabs.

Reporter plasmid construction

The plasmid pLH4254 (Figure 2.1) was digested with *NarI* at 37 °C for 1 hour followed with 20 min of 65 °C incubation to inactivate the restriction enzyme. The sequence in the 1 kb upstream of the gene of interest (GOI) and the GOI was amplified with Takara PrimeStar HS DNA Polymerase and purified with QIAquick PCR Purification Kit. The 1 kb sequence of 5' of g4535 and g4535 was amplified with JEP2849/2850. The digested pLH4254 was combined with the PCR product by using Gibson assembly® Master Mix. The assembly was transformed into NEB 5-alfa bacteria.

Knock-out plasmid construction

The plasmid pLH4252 (Figure 2.2) was digested with *NotI/ClaI* at 37 °C for 1 hour followed with 20 min of 65 °C incubation to inactivate the restriction enzyme. The sequence 1 kb 5' or 3' to the GOI was amplified with Takara PrimeStar HS DNA Polymerase and purified with QIAquick PCR Purification Kit. The 1 kb of 5' (or 3') sequence of g4535 was amplified with JEP2878/2879 (or JEP2880/2881). The digested pLH4252 was combined with the PCR products by using Gibson assembly® Master Mix. The assembly was transformed to NEB 5-alfa bacteria.

5.2 Fungal genetic modification system

RNAi feeding

It was performed as described before (AHRINGER 2006). For the strain IG1602, 25 adult transgenic worms were picked into the RNAi plate culture for 6 hour for obtaining the next generation. Other strains were seeded with synchronized L1 stage worms.

Infection

Infection plates were RNAi plates onto which spores (less than 1 month old) in solution were added. The young adult worms were washed out from RNAi bacteria and seeded on infection plates. After 2 hour of infection, worms were washed off with 50mM NaCl and rinsed another 2 times with 50mM NaCl to remove unattached spores before being transferred again to an RNAi plate. To clean the larvae, after 24 hours the worms were washed into 15 ml Falcon tube and allow to settle for 2 min. After removal the supernatant, the adult worms were transferred onto a plate spread with RNAi bacteria. After 48 hours of infection, the adult worms were separated from the larva again and analysed using the Biosort.

5.3 SapA-SapB protein interaction

These methods are used for the results section CHAPTER 4.

Formula for protocol II for fungal spore protein extraction

50 mM Tris pH 7.5

150 mM NaCL

1% NP-40

0.01% SDS

2mM EDTA

Protease inhibitors

Protocol I and other protocols used in SapA-SapB interaction experiment were described in Article 2.

Infection with JEM907 for Mass spectrometry sample preparation

Synchronized young adult N2 worms (around 5000 for each sample) from NGM plates spread with OP50 bacteria were washed out and transferred to infection plate; infection plates were OP50 topped with JEM907 spores (less than 1 month old). After 2 hour of infection, worms were washed out and seeded on other OP50 plates for 49 hour of culturing. The same method as described in 5.2 was used to clean the larva worms before collecting the infected worms.

MS protein solution preparation

The protein solution was prepared according to the APEX sample preparation protocol (REINKE *et al.* 2016). Worm pellets were directly mixed with lysis buffer by skipping the labeling procedure. Afterwards, anti-RFP beads rather than streptavidin agarose resin was used to pulling down DsRed-SapA and its interacting proteins complex. The resulting protein solutions were analysed with LC-MS-MS.

5.4 Fungal effector identification from APEX worm strains

Worm washing for APEX worm preparation

The protocol used in this experiment was adapted from Reinke's paper (REINKE *et al.* 2016), except we used a mild method to wash the worms before processing them. Instead of using centrifugation, we allowed the worms to settle on ice for 5 min to separate the worms from bacteria and fungal debris in the supernatant.

RNAi bacteria

JEM803 targeting eGFP

JEM188 targeting GFP

JEM521 *bus-8*

Worm strain

N2

rde-1 (ne300)

IG1602 *wt*; *frEx573*[*col-12p::eGFP*(partial)::SL2::*mCherry*]

IG274 *wt*;*frIs7*[*nlp-29p::GFP, col-12p::DsRed*]

Primers used

JEP2849 TAGAGGTAATCCTTCTTTGGGCACTCCGTACTTGCTGTGC

JEP2850 TCGGAGGAGGCCATGGGCCCGTAGTCAGGGGATGTTGCTT

JEP2878 CGCATGCTCCCGGCCGCCATGGCGCCATCGCGCCATGCAGGTACC

JEP2879 GTTGATGTACGGAGCACCCTAATGACGGGCGGTACGAAGCGAA

JEP2880 GCATGGACGAGCTGTACAAGTAGCACGAGAATGTCCGGGAGCTTG

JEP2881 CGTTGGGAGCTCTCCCATATGGAATAGGGGCGTCAAGTTCTGCAAC

JEP2845 ATCTTGAGCTTAGCTGATGC

JEP2847 GGTACGGCGAAGTGGTTTGG

JEP2848 CTTCGTCCTTACCAGATGAC

72 proteins presented in our and Reinke's MS data

C01B10.11	C43E11.9	F38A1.8	M28.5	Y39G10AR.8
C01B10.3	C44B7.10	F42A10.5	R05F9.6	Y46G5A.4
C02D5.4	C44E4.4	F45D11.15	R06C1.4	Y47G6A.21
C05C8.7	E01G4.6	F47B10.1	R07H5.8	Y66H1A.4
C05G5.4	F07F6.4	F49H12.5	R08E5.3	Y71H2AM.11
C09F9.2	F08D12.1	F53F1.2	R12E2.11	ZC247.1
C13B9.3	F10G8.6	F55B11.1	T06D8.10	ZC395.10
C14B9.2	F13E6.1	F57F5.1	T08A11.2	ZK1055.7
C28H8.3	F17A9.4	F57H12.6	T19B4.3	ZK632.4
C29F7.3	F20D1.3	JC8.2	T20B12.7	ZK637.2
C30C11.4	F20G2.2	K01C8.1	T25B9.9	ZK84.1
C31C9.2	F26H9.5	K01G5.5	W08E12.7	ZK909.3
C32E8.3	F27D4.1	K04G2.1	W10C8.5	ZK970.7
C35B1.5	F32D1.5	K07C5.4	Y113G7B.16	
C42D4.3	F37F2.2	K12D12.1	Y23H5B.5	

CHAPTER 6. CONCLUSION AND PERSPECTIVE

6.1 Conclusion

We have successfully adapted a PEG-mediated transformation protocol for *D. coniospora*. Together with the *ccdB* and Gibson based plasmid construction method, we established a system to genetically manipulate this fungus which serves as an important tool for host-pathogen interaction study. We identified the specific pathogenic lifestyle of *D. coniospora* based on its genomic sequence. Comparative genomic analysis revealed a list of potential fungal effectors which engage with the host immunity, for instance SapA (G3895). We further constructed a reporter strain for SapA and identified its host target SPP-5, an antimicrobial peptide. Our study focusing particularly on the pathogen provides an insight for the host-pathogen interaction between *C. elegans* and *D. coniospora*.

6.2 Perspective

Despite the successful generation of 5 *D. coniospora* transgenic strains. Nevertheless, the remaining problems such as multiple transferring during protoplasts preparation and slow growth of *D. coniospora* after transformation still need to be resolved. One of the solutions is to substitute the general medium with a medium resembling the host environment.

We show that *D. coniospora* SapA protein interacts with worm immune effector, SPP-5 *in vitro* indicating its potential role to suppress the host immunity. Due to the fact that SapA is also highly expressed at the late stage of infection, we cannot rule out the other possible functions of this protein. We could employ Mass spectrometry technique to identify other host proteins which interact with SapA *in vivo*.

The remaining question in the context of host-pathogen interaction between *C. elegans* and *D. coniospora* is how the evolutionary conserved pathogen effector interacts with its host immune effector to suppress the host immunity. For instance, we have identified SapA which is possibly adapted during the long co-evolution history with *C. elegans*. On the other hand, *C. elegans* is actively secreting various AMPs to combat with the fungus at the early stage of infection; however, of the fungal effectors used to suppress the host immunity still remain unknown. To identify other potential fungal effectors, RNA sequencing can be used on the infected sample with enriched fungal reads, such as using Fastprep a better way for releasing fungal RNA to extract the total RNA.

CHAPTER 7. REFERENCES

Ahier, A., and S. Jarriault, 2014 Simultaneous expression of multiple proteins under a single promoter in *Caenorhabditis elegans* via a versatile 2A-based toolkit. Genetics 196**:** 605-613.

Ahren, D., M. Tholander, C. Fekete, B. Rajashekar, E. Friman *et al.*, 2005 Comparison of gene expression in trap cells and vegetative hyphae of the nematophagous fungus *Monacrosporium haptotylum*. Microbiology 151**:** 789-803.

Ahringer, J., 2006 Reverse genetics, pp. in *WormBook*. The *C. elegans* Research Community ed, http://www.wormbook.org.

Aravind, L., H. Watanabe, D. J. Lipman and E. V. Koonin, 2000 Lineage-specific loss and divergence of functionally linked genes in eukaryotes. Proc Natl Acad Sci U S A 97**:** 11319-11324.

Armaleo, D., G. N. Ye, T. M. Klein, K. B. Shark, J. C. Sanford *et al.*, 1990 Biolistic Nuclear Transformation of *Saccharomyces cerevisiae* and Other Fungi. Current Genetics 17**:** 97-103.

Ballance, D. J., F. P. Buxton and G. Turner, 1983 Transformation of *Aspergillus nidulans* by the Orotidine-5'-Phosphate Decarboxylase Gene of *Neurospora crassa*. Biochemical and Biophysical Research Communications 112**:** 284-289.

Beggs, J. D., 1978 Transformation of yeast by a replicating hybrid plasmid. Nature 275**:** 104-109.

Bernard, P., P. Gabant, E. M. Bahassi and M. Couturier, 1994 Positive-Selection Vectors Using the F-Plasmid Ccdb Killer Gene. Gene 148**:** 71-74.

Bernard, P., K. E. Kezdy, L. Van Melderen, J. Steyaert, L. Wyns *et al.*, 1993 The F plasmid CcdB protein induces efficient ATP-dependent DNA cleavage by gyrase. J Mol Biol 234**:** 534-541.

Bernstein, E., A. A. Caudy, S. M. Hammond and G. J. Hannon, 2001 Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409**:** 363-366.

Billi, A. C., S. E. Fischer and J. K. Kim, 2014 Endogenous RNAi pathways in *C. elegans*. WormBook**:** 1-49.

Billmyre, R. B., S. Calo, M. Feretzaki, X. Wang and J. Heitman, 2013 RNAi function, diversity, and loss in the fungal kingdom. Chromosome Res 21**:** 561-572.

Brenner, S., 1974 The genetics of *Caenorhabditis elegans*. Genetics 77**:** 71-94.

Bundock, P., A. den Dulk-Ras, A. Beijersbergen and P. J. Hooykaas, 1995 Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. EMBO J 14**:** 3206-3214.

Case, M. E., M. Schweizer, S. R. Kushner and N. H. Giles, 1979 Efficient Transformation of *Neurospora crassa* by Utilizing Hybrid Plasmid DNA. Proceedings of the National Academy of Sciences of the United States of America 76**:** 5259-5263.

Chakraborty, B. N., and M. Kapoor, 1990 Transformation of filamentous fungi by electroporation. Nucleic Acids Res 18**:** 6737.

Chen, C. L., Y. Hu, N. D. Udeshi, T. Y. Lau, F. Wirtz-Peitz *et al.*, 2015 Proteomic mapping in live *Drosophila* tissues using an engineered ascorbate peroxidase. Proc Natl Acad Sci U S A 112**:** 12093-12098.

Chen, S. Y., X. Z. Liu and F. J. Chen, 2000 *Hirsutella minnesotensis* sp nov., a new pathogen of the soybean cyst nematode. Mycologia 92**:** 819-824.

The *C. elegans* Sequencing Consortium, T. *C. e.* S., 1998 Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science 282**:** 2012-2018.

Couillault, C., N. Pujol, J. Reboul, L. Sabatier, J. F. Guichou *et al.*, 2004 TLR-independent control of innate immunity in *Caenorhabditis elegans* by the TIR domain adaptor protein TIR-1, an ortholog of human SARM. Nat Immunol 5**:** 488-494.

Davidson, F. A., B. D. Sleeman, A. D. M. Rayner, J. W. Crawford and K. Ritz, 1996 Large-scale behavior of fungal mycelia. Mathematical and Computer Modelling 24**:** 81-87.

de Groot, M. J., P. Bundock, P. J. Hooykaas and A. G. Beijersbergen, 1998 *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. Nat Biotechnol 16**:** 839-842.

Dean, R. A., N. J. Talbot, D. J. Ebbole, M. L. Farman, T. K. Mitchell *et al.*, 2005 The genome sequence of the rice blast fungus *Magnaporthe grisea*. Nature 434**:** 980-986.

Debets, F., X. Yang and A. J. Griffiths, 1994 Vegetative incompatibility in *Neurospora*: its effect on horizontal transfer of mitochondrial plasmids and senescence in natural populations. Curr Genet 26**:** 113-119.

Dijksterhuis, J., W. Harder, U. Wyss and M. Veenhuis, 1991 Colonization and digestion of nematodes by the endoparasitic nematophagous fungus *Drechmeria coniospora*. Mycological Research 95**:** 873-878.

Dijksterhuis, J., M. Veenhuis and W. Harder, 1990 Ultrastructural study of adhesion and initial stages of infection of the nematode by conidia of *Drechmeria coniospora*. Mycological research 94**:** 1-8.

Djamei, A., K. Schipper, F. Rabe, A. Ghosh, V. Vincon *et al.*, 2011 Metabolic priming by a secreted fungal effector. Nature 478**:** 395-398.

Doehlemann, G., K. van der Linde, D. Amann, D. Schwammbach, A. Hof *et al.*, 2009 Pep1, a Secreted Effector Protein of *Ustilago maydis*, Is Required for Successful Invasion of Plant Cells. Plos Pathogens 5.

Drechsler, C., 1941 Some hyphomycetes parasitic on free-living terricolous nematodes. Phytopathology 31**:** 773-802.

Drinnenberg, I. A., G. R. Fink and D. P. Bartel, 2011 Compatibility with killer explains the rise of RNAi-deficient fungi. Science 333**:** 1592.

Dunoyer, P., G. Schott, C. Himber, D. Meyer, A. Takeda *et al.*, 2010 Small RNA Duplexes Function as Mobile Silencing Signals Between Plant Cells. Science 328**:** 912-916.

Engelmann, I., A. Griffon, L. Tichit, F. Montanana-Sanchis, G. Wang *et al.*, 2011 A comprehensive analysis of gene expression changes provoked by bacterial and fungal infection in *C. elegans*. PLoS One 6**:** e19055.

Ewbank, J. J., 2006 Signaling in the Immune Response. WormBook doi/10.1895/wormbook.1.83.1**:** 1-12.

Feinberg, E. H., and C. P. Hunter, 2003 Transport of dsRNA into cells by the transmembrane protein SID-1. Science 301**:** 1545-1547.

Felix, M. A., and C. Braendle, 2010 The natural history of *Caenorhabditis elegans*. Curr Biol 20**:** R965-969.

Fire, A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver *et al.*, 1998 Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391**:** 806-811.

Fleissner, A., A. C. Leeder, M. G. Roca, N. D. Read and N. L. Glass, 2009 Oscillatory recruitment of signaling proteins to cell tips promotes coordinated behavior during cell fusion. Proceedings of the National Academy of Sciences of the United States of America 106**:** 19393-19398.

Fleissner, A., S. Sarkar, D. J. Jacobson, M. G. Roca, N. D. Read *et al.*, 2005 The so locus is required for vegetative cell fusion and postfertilization events in *Neurospora crassa*. Eukaryotic Cell 4**:** 920-930.

Flor, H. H., 1971 Current status of the gene-for-gene concept. Annu Rev Phytopathol 9**:** 275–296.

Gibson, D. G., G. A. Benders, C. Andrews-Pfannkoch, E. A. Denisova, H. Baden-Tillson *et al.*, 2008 Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. Science 319**:** 1215-1220.

Gibson, D. G., and S. Russello, 2010 Gibson Assembly™- building a synthetic biology toolset, pp.

Gibson, D. G., L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison, 3rd *et al.*, 2009a Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods 6**:** 343-345.

Gibson, D. G., L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison, 3rd *et al.*, 2009b Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature methods 6**:** 343-345.

Graber, J. H., J. Salisbury, L. N. Hutchins and T. Blumenthal, 2007 *C. elegans* sequences that control trans-splicing and operon pre-mRNA processing. RNA 13**:** 1409-1426.

Gravato-Nobre, M. J., H. R. Nicholas, R. Nijland, D. O'Rourke, D. E. Whittington *et al.*, 2005 Multiple genes affect sensitivity of *Caenorhabditis elegans* to the bacterial pathogen *Microbacterium nematophilum*. Genetics 171**:** 1033-1045.

Gregory, P. H., 1984 The 1st Benefactors Lecture the Fungal Mycelium - an Historical-Perspective. Transactions of the British Mycological Society 82**:** 1-11.

Hartley, J. L., G. F. Temple and M. A. Brasch, 2000 DNA cloning using in vitro site-specific recombination. Genome Research 10**:** 1788-1795.

Hemetsberger, C., C. Herrberger, B. Zechmann, M. Hillmer and G. Doehlemann, 2012 The *Ustilago maydis* effector Pep1 suppresses plant immunity by inhibition of host peroxidase activity. PLoS Pathog 8**:** e1002684.

Hinnen, A., J. B. Hicks and G. R. Fink, 1978 Transformation of yeast. Proc Natl Acad Sci U S A 75**:** 1929-1933.

Hoeckendorf, A., M. Stanisak and M. Leippe, 2012 The saposin-like protein SPP-12 is an antimicrobial polypeptide in the pharyngeal neurons of *Caenorhabditis elegans* and participates in defence against a natural bacterial pathogen. The Biochemical journal 445**:** 205-212.

Houterman, P. M., B. J. Cornelissen and M. Rep, 2008 Suppression of plant resistance gene-based immunity by a fungal effector. PLoS Pathog 4**:** e1000061.

Hsueh, Y. P., P. Mahanti, F. C. Schroeder and P. W. Sternberg, 2013 Nematode-trapping fungi eavesdrop on nematode pheromones. Curr Biol 23**:** 83-86.

Hung, V., P. Zou, H. W. Rhee, N. D. Udeshi, V. Cracan *et al.*, 2014 Proteomic mapping of the human mitochondrial intermembrane space in live cells via ratiometric APEX tagging. Mol Cell 55**:** 332-341.

Irazoqui, J. E., J. M. Urbach and F. M. Ausubel, 2010 Evolution of host innate defence: insights from *Caenorhabditis elegans* and primitive invertebrates. Nat Rev Immunol 10**:** 47-58.

Jansson, H. B., 1994 Adhesion of conidia of *Drechmeria coniospora* to *Caenorhabditis elegans* wild type and mutants. J Nematol 26**:** 430-435.

Jansson, H. B., and E. Friman, 1999 Infection-related surface proteins on conidia of the nematophagous fungus *Drechmeria coniospora*. Mycological Research 103**:** 249-256.

Jansson, H. B., and B. Nordbring-Hertz, 1983 The endoparasitic nematophagous fungus *Meria coniospora* infects nematodes specifically at the chemosensory organs. Journal of General Microbiology 129**:** 1121-1126.

Jansson, H. B., A. von Hofsten and C. von Mecklenburg, 1984 Life cycle of the endoparasitic nematophagous fungus *Meria coniospora*: a light and electron microscopic study. Antonie Van Leeuwenhoek 50**:** 321-327.

Jose, A. M., G. A. Garcia and C. P. Hunter, 2011 Two classes of silencing RNAs move between *Caenorhabditis elegans* tissues. Nat Struct Mol Biol 18**:** 1184-1188.

Katayama, T., Y. Tanaka, T. Okabe, H. Nakamura, W. Fujii *et al.*, 2016 Development of a genome editing technique using the CRISPR/Cas9 system in the industrial filamentous fungus *Aspergillus oryzae*. Biotechnology Letters 38**:** 637-642.

Kim, D. H., and J. J. Ewbank, 2015 Signaling in the innate immune response, pp. 1-51 in *WormBook*, edited by T. *C. e.* R. C. ed, http://www.wormbook.org.

Kubicek, C. P., T. L. Starr and N. L. Glass, 2014 Plant cell wall-degrading enzymes and their secretion in plant-pathogenic fungi. Annu Rev Phytopathol 52**:** 427-451.

Kuersten, S., K. Lea, M. MacMorris, J. Spieth and T. Blumenthal, 1997 Relationship between 3' end formation and SL2-specific trans-splicing in polycistronic Caenorhabditis elegans pre-mRNA processing. RNA 3**:** 269-278.

Kurz, C. L., and J. J. Ewbank, 2003 *Caenorhabditis elegans*: an emerging genetic model for the study of innate immunity. Nat Rev Genet 4**:** 380-390.

Kuwano, T., C. Shirataki and Y. Itoh, 2008 Comparison between polyethylene glycol- and polyethylenimine-mediated transformation of *Aspergillus nidulans*. Current Genetics 54**:** 95-103.

Lai, Y., K. Liu, X. Zhang, K. Li, N. Wang *et al.*, 2014 Comparative genomics and transcriptomics analyses reveal divergent lifestyle features of nematode endoparasitic fungus *Hirsutella minnesotensis*. Genome biology and evolution 6**:** 3077-3093.

Larriba, E., M. D. Jaime, J. Carbonell-Caballero, A. Conesa, J. Dopazo *et al.*, 2014 Sequencing and functional analysis of the genome of a nematode egg-parasitic fungus, *Pochonia chlamydosporia*. Fungal genetics and biology : FG & B 65**:** 69-80.

Lebrigand, K., L. D. He, N. Thakur, M. J. Arguel, J. Polanowska *et al.*, 2016 Comparative Genomic Analysis of *Drechmeria coniospora* Reveals Core and Specific Genetic Requirements for Fungal Endoparasitism of Nematodes. PLoS Genet 12**:** e1006017.

Lee, K. Z., M. Kniazeva, M. Han, N. Pujol and J. J. Ewbank, 2010 The fatty acid synthase *fasn-1* acts upstream of WNK and Ste20/GCK-VI kinases to modulate antimicrobial peptide expression in *C. elegans* epidermis. Virulence 1**:** 113 - 122.

Leeder, A. C., J. Palma-Guerrero and N. L. Glass, 2011 The social network: deciphering fungal language. Nat Rev Microbiol 9**:** 440-451.

Lefrancois, S., T. May, C. Knight, D. Bourbeau and C. R. Morales, 2002 The lysosomal transport of prosaposin requires the conditional interaction of its highly conserved d domain with sphingomyelin. J Biol Chem 277**:** 17188-17199.

Li, D., and M. Wang, 2012 Construction of a bicistronic vector for the co-expression of two genes in *Caenorhabditis elegans* using a newly identified IRES. Biotechniques 52**:** 173-176.

Li, S. C., H. Kihara, S. Serizawa, Y. T. Li, A. L. Fluharty *et al.*, 1985 Activator protein required for the enzymatic hydrolysis of cerebroside sulfate. Deficiency in urine of patients affected with cerebroside sulfatase activator deficiency and identity with activators for the enzymatic hydrolysis of GM1 ganglioside and globotriaosylceramide. J Biol Chem 260**:** 1867-1871.

Lin, S., H. T. Akinbi, J. S. Breslin and T. E. Weaver, 1996 Structural requirements for targeting of surfactant protein B (SP-B) to secretory granules in vitro and in vivo. J Biol Chem 271**:** 19689-19695.

Liu, R., L. Chen, Y. Jiang, Z. Zhou and G. Zou, 2015 Efficient genome editing in filamentous fungus *Trichoderma reesei* using the CRISPR/Cas9 system. Cell Discov 1**:** 15007.

Liu, Y., S. Kuersten, T. Huang, A. Larsen, M. MacMorris *et al.*, 2003 An uncapped RNA suggests a model for *Caenorhabditis elegans* polycistronic pre-mRNA processing. RNA 9**:** 677-687.

Liu, Z., and T. L. Friesen, 2012 Polyethylene glycol (PEG)-mediated transformation in filamentous fungal pathogens. Methods Mol Biol 835**:** 365-375.

Lo Presti, L., D. Lanver, G. Schweizer, S. Tanaka, L. Liang *et al.*, 2015 Fungal effectors and plant susceptibility. Annu Rev Plant Biol 66**:** 513-545.

Lohman, U., and R. A. Sikora, 1989 Mass production of the endoparasitic fungi *Drechmeria coniospora*, *Verticillium balanoides* and *Harposporium anguillulae* in liquid culture. Nematologica 35 97–104.

Lohmann, U., R. A. Sikora and M. Höfer, 1989 Influence of Phospholipids on Growth, Sporulation and Virulence of the Endoparasitic Fungi *Drechmeria coniospora*, *Verticillium balanoides* and *Harposporium anguillulae* in Liquid Culture. Journal of Phytopathology 125**:** 139-147.

Macho, A. P., and C. Zipfel, 2014 Plant PRRs and the Activation of Innate Immune Signaling. Molecular Cell 54**:** 263-272.

Mahajan-Miklos, S., M. W. Tan, L. G. Rahme and F. M. Ausubel, 1999 Molecular mechanisms of bacterial virulence elucidated using a *Pseudomonas aeruginosa-Caenorhabditis elegans* pathogenesis model. Cell 96**:** 47-56.

Mallo, G. V., C. L. Kurz, C. Couillault, N. Pujol, S. Granjeaud *et al.*, 2002 Inducible antibacterial defense system in *C. elegans*. Curr Biol 12**:** 1209-1214.

Manning, V. A., L. K. Hardison and L. M. Ciuffetti, 2007 Ptr ToxA interacts with a chloroplast-localized protein. Molecular Plant-Microbe Interactions 20**:** 168-177.

McMullan, R., A. Anderson and S. Nurrish, 2012 Behavioral and Immune Responses to Infection Require Galphaq- RhoA Signaling in *C. elegans*. PLoS pathogens 8**:** e1002530.

Meerupati, T., K. M. Andersson, E. Friman, D. Kumar, A. Tunlid *et al.*, 2013 Genomic mechanisms accounting for the adaptation to parasitism in nematode-trapping fungi. PLoS genetics 9**:** e1003909.

Mentlak, T. A., A. Kombrink, T. Shinya, L. S. Ryder, I. Otomo *et al.*, 2012 Effector-Mediated Suppression of Chitin-Triggered Immunity by *Magnaporthe oryzae* Is Necessary for Rice Blast Disease. Plant Cell 24**:** 322-335.

Messing, J., B. Gronenborn, B. Mullerhill and P. H. Hofschneider, 1977 Filamentous Coliphage M13 as a Cloning Vehicle - Insertion of a Hindii Fragment of Lac Regulatory Region in M13 Replicative Form Invitro - (Single-Stranded-DNA Phage-Blunt End Ligation Lactose Operon Alpha-Complementation). Proceedings of the National Academy of Sciences of the United States of America 74**:** 3642-3646.

Michelle, L. P., and G. L. Hartman, 2016 Infection Mechanisms and Colonization Patterns of Fungi Associated with Soybean, Fungal Pathogenicity in *Fungal Pathogenicity*, edited by S. Sultan.

Mick, D. U., R. B. Rodrigues, R. D. Leib, C. M. Adams, A. S. Chien *et al.*, 2015 Proteomics of Primary Cilia by Proximity Labeling. Dev Cell 35**:** 497-512.

Millet, A. C., and J. J. Ewbank, 2004 Immunity in *Caenorhabditis elegans*. Curr Opin Immunol 16**:** 4-9.

Molnar, A., C. W. Melnyk, A. Bassett, T. J. Hardcastle, R. Dunn *et al.*, 2010 Small Silencing RNAs in Plants Are Mobile and Direct Epigenetic Modification in Recipient Cells. Science 328**:** 872-875.

Mueller, A. N., S. Ziemann, S. Treitschke, D. Assmann and G. Doehlemann, 2013 Compatibility in the *Ustilago maydis*-Maize Interaction Requires Inhibition of Host Cysteine Proteases by the Fungal Effector Pit2. Plos Pathogens 9.

Nakad, R., L. B. Snoek, W. Yang, S. Ellendt, F. Schneider *et al.*, 2016 Contrasting invertebrate immune defense behaviors caused by a single gene, the *Caenorhabditis elegans* neuropeptide receptor gene npr-1. BMC Genomics 17**:** 280.

Nicholas, H. R., and J. Hodgkin, 2004 Responses to infection and possible recognition strategies in the innate immune system of *Caenorhabditis elegans*. Mol Immunol 41**:** 479-493.

Nodvig, C. S., J. B. Nielsen, M. E. Kogle and U. H. Mortensen, 2015 A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi. PLoS One 10**:** e0133085.

Nordbring-Hertz, B., H.-B. Jansson and A. Tunlid, 2011 Nematophagous Fungi in *eLS*. John Wiley & Sons, Ltd.

Nowara, D., A. Gay, C. Lacomme, J. Shaw, C. Ridout *et al.*, 2010 HIGS: host-induced gene silencing in the obligate biotrophic fungal pathogen *Blumeria graminis*. Plant Cell 22**:** 3130-3141.

O'Brien, J. S., and Y. Kishimoto, 1991 Saposin proteins: structure, function, and role in human lysosomal storage disorders. FASEB J 5**:** 301-308.

O'Connell, R. J., and R. Panstruga, 2006 Tete a tete inside a plant cell: establishing compatibility between plants and biotrophic fungi and oomycetes. New Phytol 171**:** 699-718.

Okmen, B., D. W. Etalo, M. H. A. J. Joosten, H. J. Bouwmeester, R. C. H. de Vos *et al.*, 2013 Detoxification of -tomatine by *Cladosporium fulvum* is required for full virulence on tomato. New Phytologist 198**:** 1203-1214.

Olivares, C. M., and L. V. Lopez-Llorca, 2002 Fungal egg-parasites of plant-parasitic nematodes from Spanish soils. Rev Iberoam Micol 19**:** 104-110.

Ooi, S. L., J. G. Henikoff and S. Henikoff, 2010 A native chromatin purification system for epigenomic profiling in *Caenorhabditis elegans*. Nucleic Acids Res 38**:** e26.

Ottmann, C., B. Luberacki, I. Kufner, W. Koch, F. Brunner *et al.*, 2009 A common toxin fold mediates microbial attack and plant defense. Proc Natl Acad Sci U S A 106**:** 10359-10364.

Park, C. H., S. B. Chen, G. Shirsekar, B. Zhou, C. H. Khang *et al.*, 2012 The *Magnaporthe oryzae* Effector AvrPiz-t Targets the RING E3 Ubiquitin Ligase APIP6 to Suppress Pathogen-Associated Molecular Pattern-Triggered Immunity in Rice. Plant Cell 24**:** 4748-4762.

Partridge, F. A., A. W. Tearle, M. J. Gravato-Nobre, W. R. Schafer and J. Hodgkin, 2008 The *C. elegans* glycosyltransferase BUS-8 has two distinct and essential roles in epidermal morphogenesis. Dev Biol 317**:** 549-559.

Ponting, C. P., 1994 Acid sphingomyelinase possesses a domain homologous to its activator proteins: saposins B and D. Protein Sci 3**:** 359-361.

Prados Rosales, R. C., and A. Di Pietro, 2008 Vegetative hyphal fusion is not essential for plant infection by *Fusarium oxysporum*. Eukaryotic cell 7**:** 162-171.

Pujol, N., P. A. Davis and J. J. Ewbank, 2012 The Origin and Function of Anti-Fungal Peptides in *C. elegans*: Open Questions. Front Immunol 3**:** 237.

Pujol, N., O. Zugasti, D. Wong, C. Couillault, C. L. Kurz *et al.*, 2008 Anti-fungal innate immunity in *C. elegans* is enhanced by evolutionary diversification of antimicrobial peptides. PLoS Pathog 4**:** e1000105.

Qutob, D., B. Kemmerling, F. Brunner, I. Kufner, S. Engelhardt *et al.*, 2006 Phytotoxicity and innate immune responses induced by Nep1-like proteins. Plant Cell 18**:** 3721-3744.

Reinke, A. W., K. M. Balla, E. J. Bennett and E. R. Troemel, 2016 Identification of microsporidia host-exposed proteins reveals a repertoire of large paralogous gene families and rapidly evolving proteins, pp. 056788-056788. Cold Spring Harbor Labs Journals.

Rhee, H. W., P. Zou, N. D. Udeshi, J. D. Martell, V. K. Mootha *et al.*, 2013 Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. Science 339**:** 1328-1331.

Roca, M. G., L. C. Davide, L. M. Davide, M. C. Mendes-Costa, R. F. Schwan *et al.*, 2004 Conidial anastomosis fusion between *Colletotrichum* species. Mycol Res 108**:** 1320-1326.

Roca, M. G., N. D. Read and A. E. Wheals, 2005 Conidial anastomosis tubes in filamentous fungi. Fems Microbiology Letters 249**:** 191-198.

Roeder, T., M. Stanisak, C. Gelhaus, I. Bruchhaus, J. Grotzinger *et al.*, 2010 Caenopores are antimicrobial peptides in the nematode Caenorhabditis elegans instrumental in nutrition and immunity. Dev Comp Immunol 34**:** 203-209.

Romano, N., and G. Macino, 1992 Quelling: transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. Mol Microbiol 6**:** 3343-3353.

Rooney, H. C. E., J. W. van 't Klooster, R. A. L. van der Hoorn, M. H. A. J. Joosten, J. D. G. Jones *et al.*, 2005 *Cladosporium* Avr2 inhibits tomato Rcr3 protease required for Cf-2-dependent disease resistance. Science 308**:** 1783-1786.

Ruiz-Diez, B., 2002 Strategies for the transformation of filamentous fungi. Journal of Applied Microbiology 92**:** 189-195.

Sanchez-Vallet, A., R. Saleem-Batcha, A. Kombrink, G. Hansen, D. J. Valkenburg *et al.*, 2013 Fungal effector Ecp6 outcompetes host immune receptor for chitin binding through intrachain LysM dimerization. Elife 2.

Sarkies, P., and E. A. Miska, 2014 Small RNAs break out: the molecular cell biology of mobile small RNAs. Nat Rev Mol Cell Biol 15**:** 525-535.

Schaffer, U., A. Schlosser, K. M. Muller, A. Schafer, N. Katava *et al.*, 2010 SnAvi--a new tandem tag for high-affinity protein-complex purification. Nucleic Acids Res 38**:** e91.

Schmidt, S. M., H. Kuhn, C. Micali, C. Liller, M. Kwaaitaal *et al.*, 2014 Interaction of a *Blumeria graminis* f. sp hordei effector candidate with a barley ARF-GAP suggests that host vesicle trafficking is a fungal pathogenicity target. Molecular Plant Pathology 15**:** 535-549.

Schulenburg, H., and J. J. Ewbank, 2007 The genetics of pathogen avoidance in *Caenorhabditis elegans*. Mol Microbiol 66**:** 563-570.

Schulenburg, H., C. L. Kurz and J. J. Ewbank, 2004 Evolution of the innate immune system: the worm perspective. Immunol Rev 198**:** 36-58.

Shabab, M., T. Shindo, C. Gu, F. Kaschani, T. Pansuriya *et al.*, 2008 Fungal effector protein AVR2 targets diversifying defense-related Cys proteases of tomato. Plant Cell 20**:** 1169-1183.

Sijen, T., J. Fleenor, F. Simmer, K. L. Thijssen, S. Parrish *et al.*, 2001 On the role of RNA amplification in dsRNA-triggered gene silencing. Cell 107**:** 465-476.

Staskawicz, B. J., F. M. Ausubel, B. J. Baker, J. G. Ellis and J. D. Jones, 1995 Molecular genetics of plant disease resistance. Science 268**:** 661-667.

Tan, M. W., S. Mahajan-Miklos and F. M. Ausubel, 1999 Killing of *Caenorhabditis elegans* by *Pseudomonas aeruginosa* used to model mammalian bacterial pathogenesis. Proc Natl Acad Sci U S A 96**:** 715-720.

Tanaka, S., T. Brefort, N. Neidig, A. Djamei, J. Kahnt *et al.*, 2014 A secreted *Ustilago maydis* effector promotes virulence by targeting anthocyanin biosynthesis in maize. Elife 3**:** e01355.

Timmons, L., and A. Fire, 1998 Specific interference by ingested dsRNA. Nature 395**:** 854.

Tucker, S. L., and N. J. Talbot, 2001 Surface attachment and pre-penetration stage development by plant pathogenic fungi. Annu Rev Phytopathol 39**:** 385-417.

van den Burg, H. A., S. J. Harrison, M. H. Joosten, J. Vervoort and P. J. de Wit, 2006 *Cladosporium fulvum* Avr4 protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection. Mol Plant Microbe Interact 19**:** 1420-1430.

Wang, M., A. Weiberg, F. M. Lin, B. P. Thomma, H. D. Huang *et al.*, 2016 Bidirectional cross-kingdom RNAi and fungal uptake of external RNAs confer plant protection. Nat Plants 2**:** 16151.

Weiberg, A., M. Wang, F. M. Lin, H. Zhao, Z. Zhang *et al.*, 2013 Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. Science 342**:** 118-123.

Winston, W. M., C. Molodowitch and C. P. Hunter, 2002 Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1. Science 295**:** 2456-2459.

Wong, D., D. Bazopoulou, N. Pujol, N. Tavernarakis and J. J. Ewbank, 2007 Genome-wide investigation reveals pathogen-specific and shared signatures in the response of *Caenorhabditis elegans* to infection. Genome Biol 8**:** R194.

Yang, J., L. Wang, X. Ji, Y. Feng, X. Li *et al.*, 2011 Genomic and Proteomic Analyses of the Fungus *Arthrobotrys oligospora* Provide Insights into Nematode-Trap Formation. PLoS Pathog 7**:** e1002179.

Yelton, M. M., J. E. Hamer and W. E. Timberlake, 1984 Transformation of *Aspergillus nidulans* by Using a Trpc Plasmid. Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences 81**:** 1470-1474.

Yoshino, K., H. Irieda, F. Sugimoto, H. Yoshioka, T. Okuno *et al.*, 2012 Cell Death of *Nicotiana benthamiana* Is Induced by Secreted Protein NIS1 of *Colletotrichum orbiculare* and Is Suppressed by a Homologue of CgDN3. Molecular Plant-Microbe Interactions 25**:** 625-636.

Zamore, P. D., T. Tuschl, P. A. Sharp and D. P. Bartel, 2000 RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell 101**:** 25-33.

Zhang, L., Z. Zhou, Q. Guo, L. Fokkens, M. Miskei *et al.*, 2016 Insights into Adaptations to a Near-Obligate Nematode Endoparasitic Lifestyle from the Finished Genome of *Drechmeria coniospora*. Sci Rep 6**:** 23122.

Zhang, W., Z. F. Zhou, Q. N. Guo, Y. Q. Xu, M. Chen *et al.*, 2015 Culture medium and culture method for *Drechmeria coniospora* ARSEF 6962, pp. Google Patents.

Zhang, W. J., C. Pedersen, M. Kwaaitaal, P. L. Gregersen, S. M. Morch *et al.*, 2012 Interaction of barley powdery mildew effector candidate CSEP0055 with the defence protein PR17c. Molecular Plant Pathology 13**:** 1110-1119.

Zugasti, O., N. Bose, B. Squiban, J. Belougne, C. L. Kurz *et al.*, 2014 Activation of a G protein-coupled receptor by its endogenous ligand triggers the innate immune response of *Caenorhabditis elegans*. Nature immunology 15**:** 833-838.

Zugasti, O., and J. J. Ewbank, 2009 Neuroimmune regulation of antimicrobial peptide expression by a noncanonical TGF-beta signaling pathway in *Caenorhabditis elegans* epidermis. Nat Immunol 10**:** 249-256.

Zugasti, O., N. Thakur, J. Belougne, B. Squiban, C. L. Kurz *et al.*, 2016 A quantitative genome-wide RNAi screen in *C. elegans* for antifungal innate immunity genes. BMC Biol 14**:** 35.