Novel Antimicrobial Activities of *Trichoderma hamatum* GD12 Following Deletion of Heterochromatin Protein 1

Submitted by Rebecca Winsbury for the degree of

Masters by Research in Biological Sciences

May 2014

.....

Primary supervisor: Professor Chris Thornton
Secondary supervisor: Professor Murray Grant

School of Biosciences
Geoffrey Pope Building
University of Exeter
Stocker Road
Exeter
Devon
EX4 4QD

Abstract

Heterochromatin Protein 1 (HEP1) is a highly conserved, chromatin re-modelling protein involved in activation and repression of secondary metabolite producing gene clusters. In-house genome sequencing of the plant growth promoting and biocontrol fungus *Trichoderma hamatum* GD12 has shown that ~40 % of the genome is unique to GD12 compared to its closest relatives, suggesting enormous genetic potential to encode novel bioactive compounds with antimicrobial and PGP activities. It is apparent that under axenic conditions, a substantial proportion of the bioactive potential of the fungus is not expressed. We therefore hypothesised that loss of HEP1 would lead to activation of cryptic gene clusters responsible for the production of novel bioactive secondary metabolites. Identification of compounds with antimicrobial activities might benefit a growing population faced with numerous multidrug resistant microorganisms. HEP1 was inactivated in *T. hamatum* GD12 using the split-marker method of homologous recombination and Δ*ThhepA*::*hph* strains were confirmed via DIG-labelled Southern blot analysis. Phenotypic analysis revealed significantly reduced hyphal growth of *hepA* mutants compared to GD12. Confrontation assays of GD12 and three independent Δ*ThhepA*::*hph* strains against fungal pathogens revealed a change in the biocontrol activities, with a zone of inhibition surrounding mutant strains suggesting the secretion of inhibitory bioactive compound(s). Liquid chromatography-mass spectrometry was used to determine the secretome profiles of *hepA* mutants. Analysis of the data revealed a number of key features which are differentially expressed in *hepA* mutants. One such feature of particular interest is Brefeldin A, which functions as an antimicrobial agent. This project would benefit from characterisation of key features to determine their antimicrobial potentials.

Table of contents

List of Figures

List of tables

Abbreviations

4'PP	4'-phosphopantetheine
A	Adenylation
ACP	Acyl-carrier protein
AT	Acyltransferase
bp	Base pair
BSA	Bovine serum albumin
C	Condensation
c.f.u.	Colony forming units
CIA	Chloroform isoamyl alcohol
ClrD	H3K9 methyltransferase
COMPASS (CclA)	Methylation protein complex
d.p.i.	Days post inoculation
DH	Dehydratase
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
dUTP	Deoxyuridine triphosphate
E	Epimerization
EDTA	Ethylenediaminetetraaceticacid
ER	Enoyl reductase
H3K4	Histone 3 lysine 4
H3K9	Histone 3 lysine 9
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HEP1	Heterochromatin protein 1
hepA	Gene conferring HEP1
hph	Gene conferring HYG resistance
HY/YG	Split marker HYG
HYG	Hygromycin
IGV	Integrative genomics viewer
IPTG	Isopropyl-thio-β-D-galactosidase
Kb	Kilobase

KOH	Potassium hydroxide
KR	β-ketoacyl reductase
KS	Ketoacyl synthase
LaeA	Global regulator protein
LC-MS	Liquid chromatography - mass spectrometry
LF	Left flank
MAPK	Mitogen activating protein kinase
Mb	Megabase
MEA	Malt extract agar
mM	Millimolar
MT	Methyltransferase
NRPS	Nonribosomal peptide synthatase
ORF	Open reading frame
PCP	Peptidyl carrier protein
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PkaA	Protein kinase
PKS	Polyketide synthase
Pol II	RNA polymerase II
QTOF	Quantitative time of flight
r.p.m.	Rotations per minute
RF	Right flank
RNA	Ribonucleic acid
RT	Retention time
SAT	Starter ACP transacylase
SDS	Sodium dodecyl sulfate
SM(S)	Secondary Metabolite(s)
TE	Thioesterase
VeA	Velvet protein A
VelB	Velvet-like protein B
w/v	Weight by volume
w/w	Weight by weight
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

Acknowledgements

For invaluable guidance and support throughout the course of this project I would like to thank my primary supervisor Professor Chris Thornton. Thank you for your supervision, I am truly grateful.

For encouragement, support and also for providing a scholarship for this project, I would like to thank my secondary supervisor Professor Murray Grant. Thank you for all of your generous support and guidance during my time in lab301, and for all of the opportunities you have provided me with, I am truly grateful.

I would like to say a huge thank you to Dr. Marta De-Torres. Thank you for always taking the time to help me, and for your kind words of support. You really are an inspiration.

A thank you to all of the 'Scallywags'. Kate, thank you for guiding me through this project and for making me laugh with 'dad' jokes. Odette, thank you for your support and for making every day in the lab a joy, and also for being clumsier than me! And Genna, Rosie, Trupti and Marwan, you made every day in the lab enjoyable.

Thank you to Hannah and Christine for your superb help on the LC-MS aspect of this project! Truly invaluable support and an excellent ending to this thesis.

To Roger, thank you for putting up with my 'constant state of stress' and 'OCD', although I'd still go with 'particular'.

I would like to thank family for always believing in me, and supporting me unconditionally.

Finally, I would like to thank members of lab 301, both new and old. The success of this project was a result of help and support from every individual. Thank you.

1. Introduction

Trichoderma species are ubiquitous soil saprotrophs that are renowned for their prolific secretion of secondary metabolites (SMs)[1], low-molecular-mass compounds which, unlike primary metabolites, are not essential for the growth and development of the organism producing them. Plants and microorganisms produce SMs as a survival mechanism and humans have been able to utilise these compounds for medical use due to their pharmaceutical and toxicological properties[2]. With the global population expected to exceed 9 billion by 2050[3] and antibiotic resistance becoming more prominent[4], the ability to identify microorganisms that produce novel compounds with active biocontrol capabilities on a large scale becomes increasingly important.

Substantial portions of microbial genomes are dedicated to the production of SMs[5,6], yet under standard laboratory conditions it is clear that the overwhelming majority of this biosynthetic potential is not expressed[5–7].

Secondary metabolites are derived from complex gene clusters[8] which are often located within the subtelomeric region of chromosomes[9,10]. This project set out to activate such a cryptic gene cluster in the beneficial rhizosphere fungus *T. hamatum* GD12 in an attempt to identify novel compounds for use as antimicrobial agents.

1.1 *Trichoderma hamatum* GD12 is unique

Trichoderma spp., are members of the largest group of fungi, the Ascomycota. Some strains have been shown to elicit plant-growth-promotion (PGP) through secretion of PGP compounds[11,12], whereas others display biocontrol against a broad range of pathogens through a variety of mechanisms. For example, activation of induced systemic resistance (ISR) in plants by *Trichoderma* spp., has been shown to elicit biocontrol against the bacterial pathogen *Pseudomonas syringae*[13] and the fungal pathogen *Botrytis cinerea*[14]. Competition for nutrients is the mode of action used by many *Trichoderma* spp. such as biocontrol of *Fusarium* wilt disease of tomatoes by *Trichoderma asperellum*[15]. Many *Trichoderma* spp. produce cell wall-degrading enzymes such as chitinase that medicate biocontrol during physical interactions with pathogens, such as during antagonism of the devastating white mould fungus, *Sclerotinia sclerotiorum*[16]. *Trichoderma* spp. also produce secondary metabolites which have been shown to elicit PGP activity[17].

The *T. hamatum* GD12 genome was sequenced in-house and a draft genome assembled from 12 million pairs of Illumina GA2 paired-end 73-bp reads using Velvet 1.1.04. BLASTP analysis of the 38.2 Mb whole genome shotgun sequence, deposited at DDBJ/EMBL/GenBank under the accession ANCB00000000[18], against four other *Trichoderma* spp. revealed a conserved 'core' proteome of 3620 proteins. The close relative *T. atroviride* shares an additional 2096 predicted proteins with GD12 supporting the hypothesis that divergence of species occurred very recently on an evolutionary scale. A large portion of the predicted proteome is unique to *T. hamatum* GD12, comprising 4658 predicted proteins, some with potential bioactive capabilities (Figure 1A).

Prediction of the *T. hamatum* GD12 secretome was carried out via SignalP 4.0[19] analysis against four other *Trichoderma* species revealing a secretome of 327 core predicted secretion proteins found unanimously across all five of the *Trichoderma* spp. analysed (Figure 1B). Interestingly, the number of predicted secretome proteins unique to *T. hamatum* is greater than that of the predicted 'core' secretome, at 370 predicted proteins. Evidence of recent species divergence between *T. hamatum* and *T. atroviride* is present again within the predicted secretome, with 164 secretion proteins predicted to be shared between the two fungi, the highest of all *Trichoderma* pairwise comparisons.

From these findings we hypothesised that, based on the large number of predicted and secreted proteins which are unique to GD12, it is highly probably that there are cryptic gene clusters present within *T. hamatum* GD12 that have potential to encode an abundance of novel bioactive secondary metabolites.

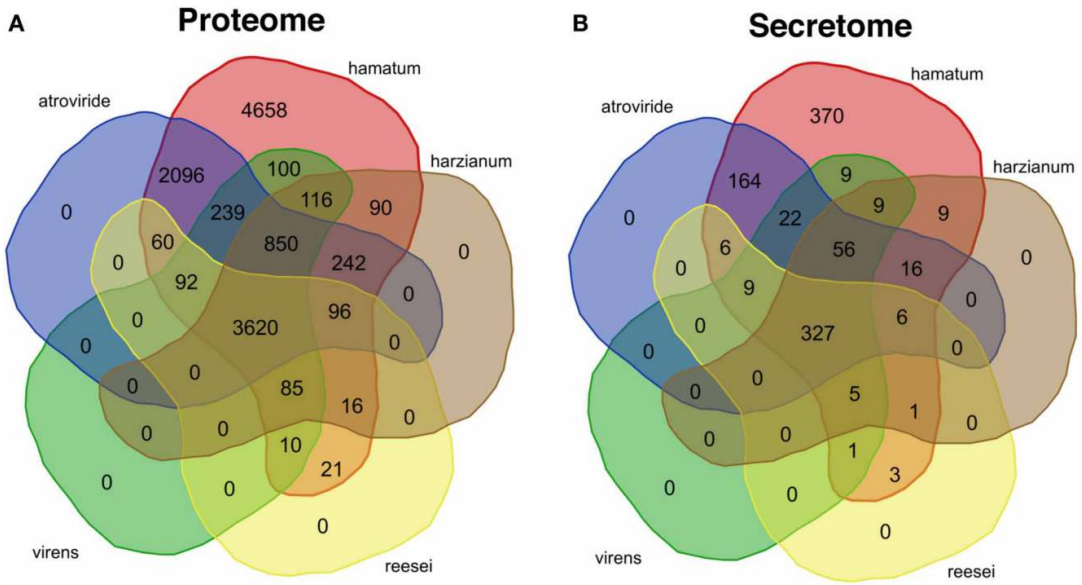


Figure 1 | Venn diagrams showing the conservation of the predicted proteome and secretome of *Trichoderma hamatum* GD12 compared with other *Trichoderma* spp. [Taken from Studholme *et al.* 2013[18]] The GD12 genome was sequenced in-house using Illumina GA2 paired-end 73-bp reads using Velvet 1.1.04. BLASTP and SignalP were used to analyse the conservation of the predicted proteome and secretome of GD12 compared with 4 other *Trichoderma* spp. **(A)** *T. hamatum* GD12 shares a core proteome of 3620 predicted proteins with *T. atroviride, T. virens, T. reesei* and *T. harzianum*. *T. hamatum* GD12 also has 4658 unique predicted proteins and shares close homogly with *T. atroviride*. **(B)** A predicted secretome analysis from SignalP identified 327 core secreted proteins common to all *Trichoderma* spp. and 370 which were unique to *T. hamatum* GD12 alone.

1.2 Synthesis of secondary metabolites through biosynthetic pathways

Secondary metabolites are derived from biosynthetic genes typically found in clusters[20] within the sub-telomeric region of chromosomes, although there are exceptions[21]. Biosynthetic genes encode large multidomain, multimodular enzymes (Figure 3) which synthesize polyketides (PKS), such as the mycotoxin aflatoxin B1, or non-ribosomal peptides (NRPS) such as the common antibiotic penicillin G (Figure 2). Although the majority of SMs are derived from one of these two pathways, there are exceptions. Some SMs, such as the toxin coronatine, produced by the bacterial plant pathogen *Pseudomonas syringae* during host infection, are derived from a hybrid PKS-NRPS[22,23]. The SM ergotamine is derived from NRPS but requires a tryptophan dimethylallytransferase for synthesis[24] and the plant hormones gibberellins, which are also produced by a range of fungi, are derived independently of PKS and NRPS but require a terpene cyclase for synthesis[25].

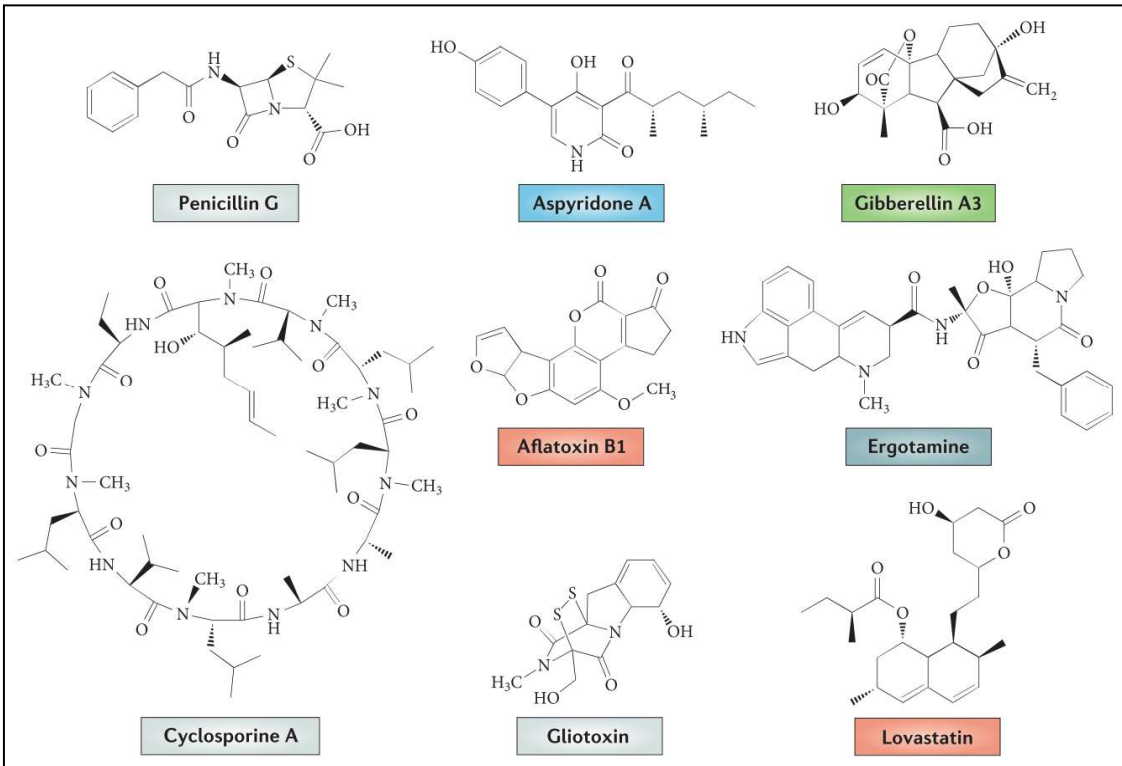


Figure 2 | Examples of fungal secondary metabolites. [Taken from Brakhage, 2013[29]] Fungal secondary metabolites are produced through multimodular, multidomain biosynthetic pathways. Penicillin G, cyclosporine A and gliotoxin are all derivatives of non-ribosomal peptide synthetases (NRPS - represented in light grey). Aflatoxin B1 and lovastatin are derived from polyketide synthases (PKS – represented in red). NRPS and PKS constitute the majority of secondary metabolites. Others include aspyridone A which is derived from the hybrid pathway PKS-NRPS (represented in blue), ergotamine which is derived from the NRPS pathway but requires a tryptophan dimethylallyltransferase for synthesis (shown in dark grey), and finally gibberellin A3, plant hormones which are also produced by some fungi and derived independently from both PKS and NRPS pathways but require a terpene cyclase for synthesis (shown in green).

Synthesis of SMs begins with malonyl and amino acid building blocks for PKS and NRPS respectively, or derivatives thereof[26–28], which are passed along a series of modules, each of which is responsible for one discrete elongation step (Figure 3). For NRPS synthesis, three minimal domains are required; an **adenylation domain** (A: activation of amino acid building block), a **peptidyl carrier protein** (PCP, also known as thiolation domain: binds cofactor 4'PP to which an activated amino acid covalently attaches), and a **condensation domain** (C: catalyses peptide bond formation). Additional extensions and modifications may include a methyltransferase (MT: addition of a methyl group), a β-ketoacyl reductase (KR: reduction of a ketoacyl group), and/or an epimerization (E: changing of one asymmetric centre in a compound). Similarly to NRPS, PKS also requires a minimum of three domains; an **acyltransferase domain** (AT: extender unit selection and transfer), an **acyl-carrier protein** (ACP: extender unit loading), and a **ketoacyl synthase domain** (KS: decarboxylative condensation of extender unit with an acyl thioester). The resulting product is a β-ketothioester which may undergo additional elongation and modification via a β-ketoacyl reductase domain, a dehydratase domain (DH: loss of H_2O), an enoyl reductase domain (ER: reduces a β-double bond to a single bond), and/or a methyltransferase domain.

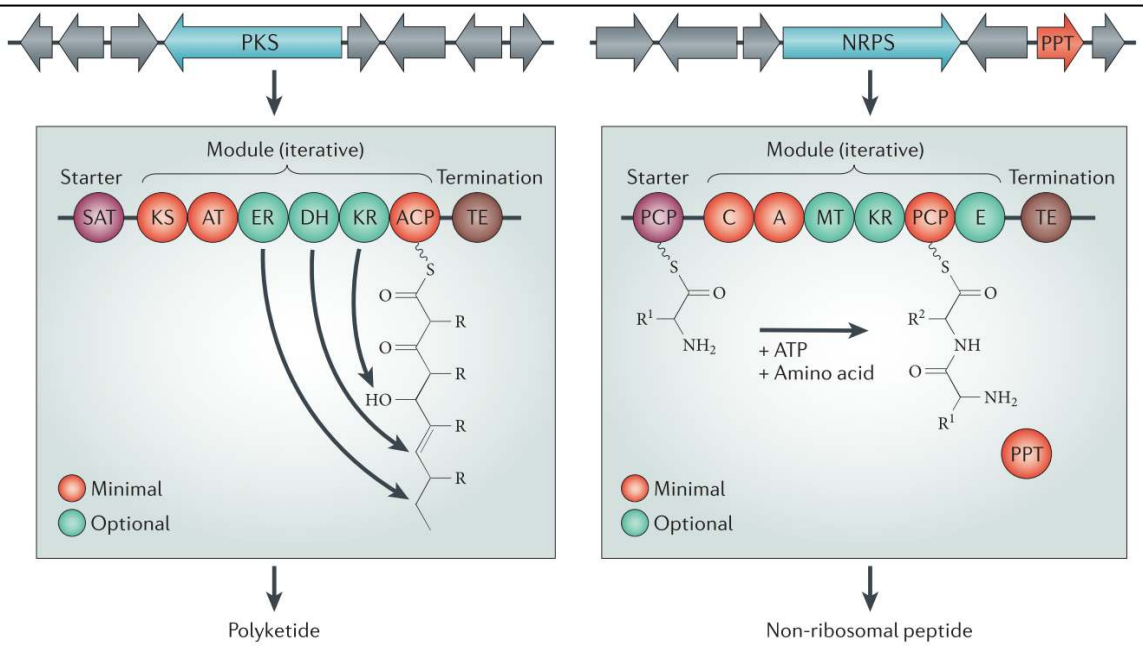


Figure 3 | Synthesis of secondary metabolites through gene clusters. [Taken from Brakhage, 2013[29]] Secondary metabolites are derived from biosynthetic pathways which contain multimodular, multidomain enzymes, each of which is responsible for one discrete elongation step. **Polyketide** synthase domains (left) begin with a malonyl starter unit and require a minimum of three domains for synthesis; a ketoacyl synthase (KS) domain, an acyltransferase (AT) domain and an acyl-carrier protein (ACP) domain. Optional elongation steps may include; an enoyl reductase (ER) domain, a dehydratase (DH) domain and a β-ketoacyl reductase (KR) domain. These enzymes are flanked by a starter ACP transacylase (SAT) domain and a termination domain (TE). **Non-ribosomal peptide synthetases** begin with an amino acid starter unit and also contain a minimum of three domains for synthesis; a condensation (C) domains, an adenylation (A) domain and a peptidyl carrier protein (PCP) domain. Additional elongation steps may include; a methyltransferase (MT) domain, a β-ketoacyl reductase (KR) domain and an epimerization (E) domain. These enzymes are flanked by a PCP domain, which binds the cofactor 4'-phosphopantetheine (4'PP), to which the activated amino acid starter unit is covalently attached and a termination domain.

1.3 Fungal secondary metabolism

Chromatin exists in two forms: euchromatin, the 'on' state when genes are actively being transcribed, and the condensed state - heterochromatin, when the gene cluster is repressed. It is not yet fully understood how this pathway functions, but a proposed mechanism is illustrated in Figure 4 [30].

Methylation residues on lysine 4 of histone 3 (H3K4-CH_3) and acetylation residues are commonly associated with active gene transcription in euchromatin. Unknown silencing factors are thought to initiate histone deacetylase (HDAC) to remove acetylation residues from histones[31], and the addition of a methyl group to histone 3 lysine 9 is achieved via H3K9 methyltransferase (ClrD), subsequently creating the binding site for Heterochromatin Protein 1 (HEP1)[32]. The addition of HEP1 to the complex causes chromatin to condense into heterochromatin, effectively silencing the gene cluster. Gene clusters remain silent until they are reactivated when unknown environmental stimuli are translated into signal cascades which are able to act both dependently and independently through the LaeA containing Velvet complex[33]. The removal of methylation from H3K9 by the COMPASS complex[34] and subsequent acetylation via histone acetyltransferase (HAT)[35] initiates gene transcription through RNA polymerase II (Pol II)[36].

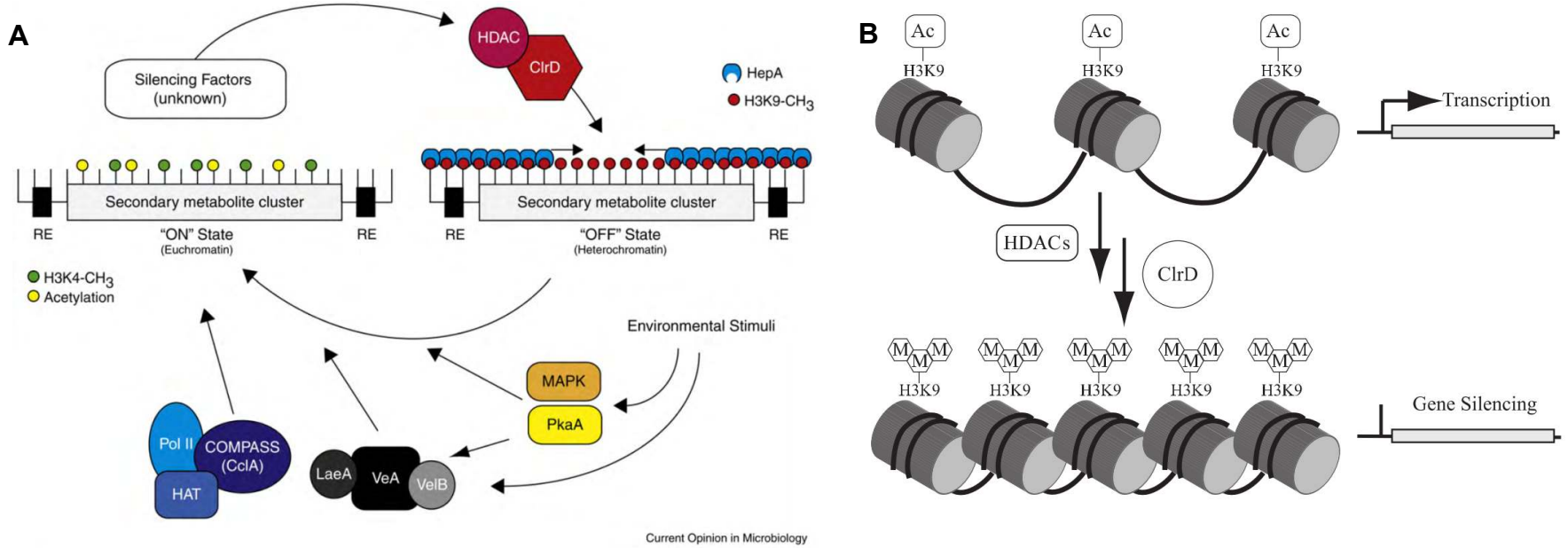


Figure 4 | A proposed model for chromatin regulation of secondary metabolite gene clusters. [Taken from Palmer *et al.* 2010[30] (**A**) and Palmer *et al* 2008[37] (**B**)]. It is thought that unknown silencing factors initiate histone deacetylase to remove acetyl groups from chromatin, which are commonly associated with active gene transcription, euchromatin. This is turn allows methylation of histone 3 lysine 9 (H3K9) by ClrD, creating the binding site for heterochromatin protein 1, HEP1. The addition of HEP1 to the complex causes to chromatin to condense into the silent state of heterochromatin, repressing the gene cluster. Initiation of active transcription occurs when environmental stimuli are transcribed into signal cascades which act both dependently and independently of the LaeA containing Velvet Complex. Histone acetyltransferase (HAT) allows methylation of H3K4 and the COMPASS complex allows active gene transcription through RNA polymerase II (Pol II).

1.4 Heterochromatin Protein 1 (HEP1)

Heterochromatin protein 1 (HEP1) is a highly conserved eukaryotic protein first identified in *Drosophila melanogaster* as a dominant suppressor of position-effect variegation[38] – translocation of euchromatic genes to the vicinity of pericentric heterochromatin where they acquire a variegated pattern of expression. Since this initial discovery, orthologs have been found in a broad range of eukaryotic organisms, with many carrying multiple copies (Figure 5)[39]. Genomes of the fission yeast *Schizosaccharomyces pombe* and the red bread mould *Neurospora crassa* each contain one HP1 homolog, whereas the soil-living amoeba *Dictyostelium discoideum* has two. Some animal species are known to have up to five HP1 orthologs within their genomes with 50% amino-acid sequence identity between mammalian HP1 and the homolog found within *Drosophila*[40].

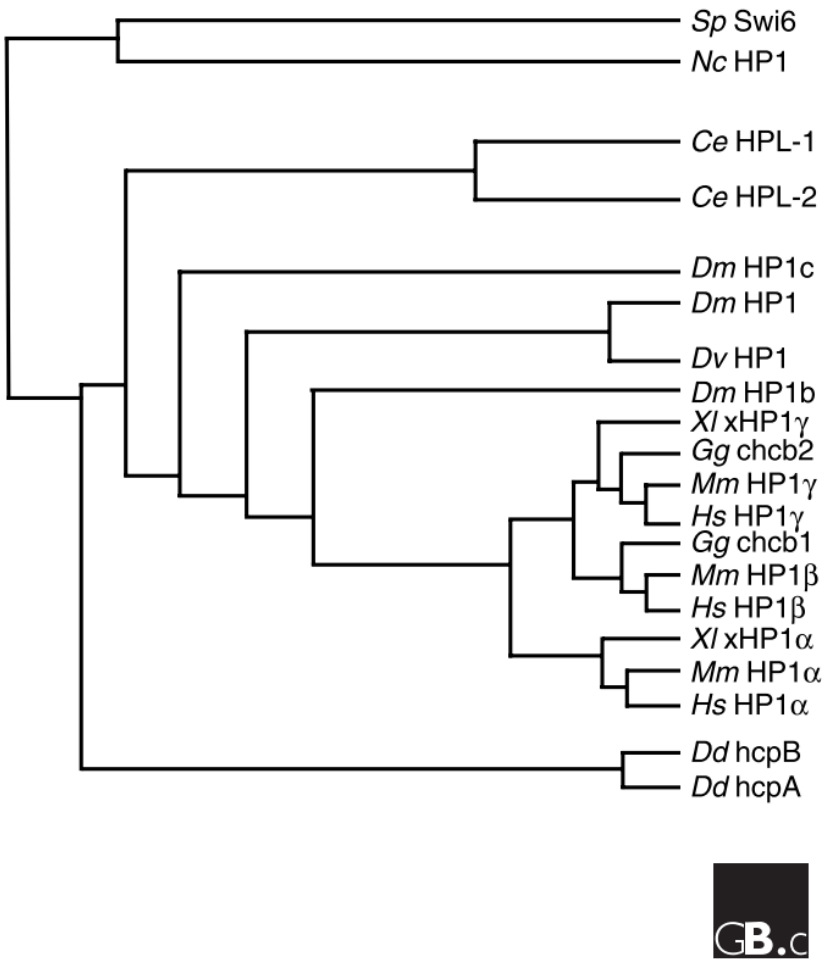


Figure 5 | Phylogenetic tree showing Heterochromatin Protein 1 homology. [Taken from Lomberk *et al* 2006[39]]. Heterochromatin protein 1 is a highly conserved eukaryotic protein with roles in activation and silencing of gene clusters by chromatin remodelling. Species shown include the fission yeast *Schizosaccharomyces pombe* (Sp), the red bread mould *Neurospora crassa* (Nc), the soil nematode *Caenorhabditis elegans* (Ce), the common fruit fly *Drosophila melanogaster* (Dm) and its close relative *Drosophila virilise* (Dv), the African claw-toed frog *Xenopus laevis* (Xl), *Gallus gallus* more commonly known as the red junglefowl (Gg), the house mouse *Mus musculus* (Mm), the human species *Homo sapiens* (Hs) and the soil-living amoeba *Dictyostelium discoideum* (Dd).

Structurally, HEP1 consists of two highly conserved domains, the chromo domain and the chromo-shadow domain, which are connected via a variable linker region (Figure 6)[39]. The amino terminal half of the protein, the chromo domain, is responsible for gene-silencing by altering the structure of chromatin to produce heterochromatin. This is achieved by three conserved aromatic residues, identified as tyrosine-24 (Tyr(Y)), tryptophan-45 (Trp(W)) and Tyr-48 within *Drosophila*, which form a three walled aromatic cage creating a hydrophobic pocket which allows the chromo domain to dock with methylation residues on di- and trimethylated H3K9[41,42]. The carboxy-terminal half of the protein, the chromo-shadow domain, is responsible for homo- and heterodimerization and interaction with other chromatin associated molecules[32,43]. The linker or hinge-region which separates the chromo domain from the chromo-shadow domain contains the least conserved amino acid sequence between HEP1 orthologs, and is thought to be flexible and exposed to the surface[44].

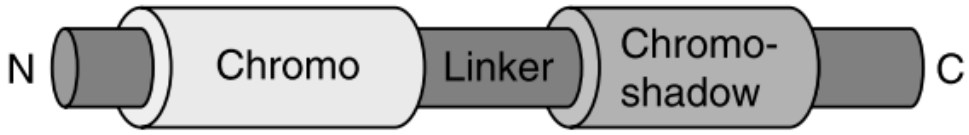


Figure 6 | The conserved linear structure of Heterochromatin Protein 1. [Taken from Lomberk *et al* 2006[39]]. The highly conserved structure of heterochromatin protein 1 consists of an amino (N)-terminal chromo domain which binds to chromatin altering its structure, and the carboxy (C)-terminal chromo-shadow domain which is responsible for binding to other chromatin associated molecules. These two highly conserved domains are connected by a more variable linker region.

1.5 Project aims

Trichoderma spp. are renowned for being prolific producers of bioactive secondary metabolites which have been utilised in a range of applications, such as medicine and chemical manufacturing. Under axenic laboratory conditions it appears that an overwhelming proportion of the bioactive potential is not expressed. Activation of such cryptic gene clusters may reveal novel secondary metabolites with bioactive properties. Analysis of the genomic gene clusters involved in secondary metabolite production suggest a 'mosaic'-type pathway, where regulatory proteins, such as HEP1, play a key role in chromatin re-modelling.

The objective of this project was to identify a HEP1 homolog within *T. hamatum* GD12 and to investigate the antimicrobial activities of HEP1-deficient mutants.

2. Materials and methods

2.1 Growth and maintenance of strains

Long term growth and maintenance of all strains and pathogens used throughout this project (Table 1) was carried out on malt extract agar (MEA: 2 % [w/v] malt extract, 2 % [w/v] agar) and grown for experimentation on potato dextrose agar (PDA: 2.4% [w/v] potato dextrose, 2 % [w/v] agar) unless stated otherwise. During active growth, fungal and yeast strains were incubated at 26 °C with a 24 h light cycle consisting of 16 h light and 8 h dark.

Table 1 | Strains and isolates used within this project.

KD: Katherine Denby, Life Sciences, University of Warwick. SB: Steve Bates, School of Biosciences, University of Exeter. CBS: Centraalbureau voor Schimmelcultures, Utrecht, The Netherlands. JW: Jon West, Rothamsted Research, Harpenden, Hertfordshire.

Organism	Isolate number	Source
Botrytis cinerea	R2	KD
Candida albicans	SC5314	SB
Candidia tropicalis	1920	CBS
Filobasidiella (Cryptococcus) neoformans	10490	CBS
Fusarium oxysporum f.sp. lycopersici	167.3	CBS
Geotrichum candidum	115.23	CBS
Pythium ultimum var. *ultimum*	656.68	CBS
Rhizoctonia solani		
Sclerotinia minor		
Sclerotinia sclerotiorum	BFS	JW
Sclerotinia sclerotiorum	GFR1	JW
Sclerotinia sclerotiorum	GFR11	JW
Sclerotinia sclerotiorum	M448	JW
Trichosporon asahii	892	CBS
Trichosporon asteroides	6183	CBS
Trichosporon inkin	7630	CBS

2.2 Fungal genomic DNA extraction

T.hamatum GD12 mycelium was obtained by inoculating 100 mL potato dextrose broth (PDB: 2.4% [w/v] potato dextrose) with 4 x 5 mm diameter plugs of actively growing *T. hamatum* GD12 mycelia, taken from the leading edge of cultures 2 days post inoculation (d.p.i.). Cultures were incubated at 26 °C for 72 h at 125 rotations per minute (rpm) before filtering through sterile miracloth, washing with sterile deionised water and flash-freezing in liquid nitrogen. Mycelia was ground to a fine powder using a sterile pestle and mortar, which had been chilled using liquid nitrogen. The powder was transferred to a 2 mL microfuge tube containing 1 mL SDS-buffer (1 % [w/v] SDS, 0.025 M EDTA (pH 8.0), 0.25 M NaCl, 0.2 M Tris-HCl (pH 8.0)) and incubated at 65 °C for 30 min. Cellular debris was removed by centrifugation at 14,000 rpm for 10 min at room temperature and the supernatant decanted into a fresh sterile 1.5 mL microfuge tube. For nucleic acid purification, 800 µL phenol (pH 8.0) was added to each tube, the tubes vortexed and residual debris removed by centrifugation as previously described. To further purify the aqueous phase containing nucleic acid, two further extractions were carried out with 800 µL phenol:chloroform [1:1] and CIA, respectively. Nucleic acid was precipitated from the supernatant by adding 0.6 vol. of ice-cold isopropanol, vortexing for an even distribution and incubating at -20 °C for 30 min. Nucleic acids were harvested by centrifugation at 14,000 rpm for 10 min at 4°C, the supernatant removed and the pellet washed with 500 µL ice-cold 70 % [v/v] ethanol. Samples were pelleted for 5 min at 14,000 rpm, room temperature, residual ethanol removed with a pipette and the pellet dried before re-suspension in 30 µL sterile milliQ water. RNA was removed with RNase and the concentration of DNA determined by Nanodrop

spectrophotometer (Thermo Scientific) and gel electrophoresis.

2.3 Creation of Δ*ThhepA*::*hph* mutants

2.3.1 Creation of *hepA* knockout cassette

The split marker method of homologous recombination was used to replace the *hepA* gene of *T. hamatum* with the hygromycin resistance conferring *hph* gene from *Escherichia coli* under a *Neurospora crassa* promoter. The split marker cassettes were created by a series of PCR reactions outlined in Appendices 6.1 and 6.2. Each master mix consisted of 9.5 µL GoTaq® Green Master Mix (Promega), 1 µL forward primer (10 pM), 1 µL reverse primer (10 pM), 50 ng template DNA and sterile milliQ water to a final volume of 25 µL. PCR reactions were analysed by gel electrophoresis on an agrose gel (0.8% [w/v] in TAE) containing 0.5 µg mL ethidium bromide to visualise DNA. PCR products were run alongside a GeneRuler 1kB ladder (Fermentas). QIAquick Gel Extraction Kit (Qiagen) was used to purify PCR products from agarose gels according to the manufacturer's protocol and the final concentration of purified products were determined by gel electrophoresis (Figure 9).

2.3.2 Protoplast transformation of *Trichoderma hamatum* GD12

Conidia of *T.hamatum* GD12 were harvested from V8 agar plates (V8: 20 % [v/v] V8 juice, 1 % [w/v] D-glucose), 2 % [w/v] agar, to volume with milliQ water) 7 d.p.i by agitation in sterile milliQ water. Twenty mL PDB was inoculated with 10^6 c.f.u mL^{-1} and incubated at 26 °C for 48 h, static. Mycelia were harvested via filtration through sterile miracloth and residual PDB removed by washing with sterile milliQ H_2O. Fungal biomass was weighed to 0.6g and incubated in 2.4 mL filter sterilised (0.22 µm) enzyme solution containing 1.2 mg chitinase

(Sigma), 1 mg lyticase (Sigma) and 44 mg cellulose (Sigma) in mannitol osmoticum ((50 mM $CaCl_2$, 0.5 M Mannitol) adjusted to pH 5.5 with KOH)) at room temperature and shaken at 225 rpm. After 25 min, the protoplast concentration was determined by counting under a haemocytometer (10^7 protoplasts mL^{-1} desired), with the incubation period preceding no longer than 45 min. Fungal debris was removed from the protoplast mixture by filtering through sterile miracloth and the protoplasts harvested by centrifugation at 5,500 rpm for 5 min at 4 °C. The pellet was gently re-suspended in 300 µL filter sterilised mannitol osmoticum. The centrifugation and re-suspension step was repeated two times to remove any residual enzyme mixture, with the protoplasts finally being re-suspended in 240 µL filter sterilised mannitol osmoticum. Five µg of each *hepA* LF + HY and *hepA* RF + YG purified PCR products (no more than 40 µL total) were added to the protoplast suspension then the mixture incubated on ice for 20 min before 130 µL of PEG solution (40 % [w/v] PEG 4000 in mannitol osmoticum) was added and mixed by inversion. A further 130 µL of PEG solution was added to the mixture, again mixed by inversion, and incubated at room temperature for 30 min. The protoplast suspension was gently mixed with 150 mL PDA + sucrose agar (PDA with 0.8 M sucrose) at 42 °C which was poured into five x 9 cm petri dishes to solidify. Plates were incubated in the dark at 26 °C for 24 h before a PDA + 600 µg mL^{-1} hygromycin overlay was applied as a selection layer to each plate. Plates were returned to the dark and checked daily for putative transformants which were isolated from the surface of the overlay layer and sub-cultured on to PDA + 600 µg ml^{-1} hygromycin. To ensure stability of the *hph* gene and also to ensure that each putative transformant was selected from an individually transformed protoplast, single spore isolation was carried out by growing each strain on V8 agar and

harvesting conidia 7 d.p.i. Conidia were diluted to 10^2 c.f.u^{-1} and grown on PDA + 600 μg^{-1} at 26 °C until individual transformants emerged. These were selected and sub-cultured on to V8 agar and the single spore isolation process was repeated once more.

2.4 DIG-Southern blot analysis

2.4.1 Creation of DIG-labelled left flank probe

To create the DIG-labelled probe, PCR of the left flanking region was carried out using 40 µL buffer HF 5X (Promega), 20 µL DIG-labelled nucleotides (Roche), 4 µL hepA_LF_LP, 4 µL hepA_LF_RP, 2 µL Phusion® Taq DNA Polymerase (New England BioLabs, NEB), 50 ng of previously purified LF product and sterile milliQ water to a final volume of 200 µL. The master mix was divided in to four x 50 µL aliquots and the PCR cycle run with an annealing temperature of 55 °C and a 30 s extension time. PCR products were analysed on a 0.8 % TAE agarose gel and the product purified as described above with the final product being eluted in 20 µL milliQ water before being added to 20 mL pre-warmed (to eliminate precipitation) Southern Hybridization buffer ($NaPO_4$ - pH 7 (0.5 M), 7 % SDS).

2.4.2 Confirmation of Δ*ThhepA*::*hph* strains

Genomic DNA was extracted from the putative transformants as described previously, and 20 µg of each putative transformant, plus a GD12 control, were digested using 3 µL *Stu*I restriction enzyme (NEB), 5 µL buffer 10X, 0.5 µL BSA and sterile milliQ water to a final volume of 50 µL. Restriction digests were incubated overnight at 37 °C and the following day were run on a 0.8 % TAE agarose gel. The gel was placed well-side down in a trough and depuration was carried out by shaking for 15 min in 50 mL 0.25 M HCl. The HCl was then replaced with 50 mL 0.4 M NaOH for a further 15 min to allow neutralisation. To transfer the digested DNA to a membrane, a blot was performed by filling a

large trough with 0.4 M NaOH and placing a piece of Perspex® plastic over the trough as a bridge. A large strip of Whatman® paper soaked in 0.4 M NaOH was draped over the bridge with each end in the 0.4 M NaOH solution to act as a wick. The gel was placed in the centre of the Whatman® paper wick with wells facing up. A piece of Amersham Hybond-NX membrane (GE Healthcare) cut to the size of the gel was placed on top of the gel using forceps (the membrane was not moved once touching the gel), followed my two pieces of Whatman paper cut to the same size as the gel and finally a stack of paper towels. Another piece of Perspex was placed on top of the paper towels and pressure applied to the stack from the top. Seran^{TM} wrap was placed down each side of the bridge to enhance the capillary effect and prevent precipitation of the 0.4 M solution and the blot was left overnight. The paper towels and Whatman® paper squares were discarded and the membrane transferred to a hybridisation tube along with 50 mL Southern Hybridization Buffer and incubated at 62 °C for 30 min. Meanwhile, the probe was boiled in a 100 °C water bath for 10 min. The Southern Hybridization buffer in the hybridization tube was discarded and replaced with the boiled probe and the membrane was then left to incubate at 62°C overnight. The probe was removed and the membrane washed twice for 15 min at 62 °C with 20 mL Southern Wash Buffer ($NaPO_4$ – pH 7 (0.1 M), 1 % [w/v] SDS) in the hybridization tube. The membrane was transferred to a trough and washed for 5 min in 20 mL DIG-wash buffer (0.3 % Tween 20, DIG Buffer 1 (maleic acid (0.1 M), NaCl (0.15 M) adjusted to pH 7 with NaOH)) at room temperature with agitation. The DIG-wash buffer was then removed and replaced with 25 mL DIG Buffer 2 (1 % [w/v] semi-skimmed milk powder in DIG Buffer 1) and the membrane incubated for 30 min with agitation. After 30 min, the Blocking solution was replaced with 20 mL Antibody solution (anti-DIG AB

(Roche) as a 1:10000 dilution) in DIG Buffer 2) and incubated for a further 30 min at room temperature with agitation. The membrane was then washed twice with 20 mL DIG-wash buffer for 15 minutes at room temperature with agitation, and then equilibrated for 5 min in 20 mL DIG-buffer 3 (Tris-HCl (0.1 M), NaCl (0.1 M) $MgCl_2$ (50 mM); adjusted to pH 9.5 with HCl). The membrane was removed from the trough using forceps and placed in a plastic envelope with 1 mL CDP-Star (Roche) solution pipette on the surface of the membrane. The plastic envelope was closed and bubbles removed, to ensure the CDP-star solution covered the entire surface of the membrane, and incubated at room temperature for 5 min. CDP-star was then thoroughly drained from the membrane on to paper towel and the membrane sealed in a fresh plastic envelope, placed into a film cassette and incubated for 15 min at 37 °C. In a dark room, an X-ray film (Fujifilm) was exposed to the membrane and the cassette closed and incubated at room temperature for 1 min before the film was developed.

2.5 Phenotypic Analysis

2.5.1 Growth curves

T. hamatum strains were grown on PDA in 15 cm square petri dishes and incubated at 26 °C in both light and dark conditions. Plates were scanned daily over a 14 day period using an Epson Perfection V750 Pro scanner, and growth of mycelia recorded as mm^2 using imageJ.

2.5.2 Confrontation assays

To determine the biocontrol effects loss of HEP1 has on various pathogens in comparison to the wild-type strain, 5 mm plugs of mycelia from *T. hamatum* strains and fungal pathogens (Table 1) were taken from the leading edge of cultures 2 d.p.i and placed at opposite sides of 9 cm petri dishes containing PDA. Yeast pathogens were sub-cultured on to PDA plates containing *T. hamatum* strains 3 d.p.i using a sterile inoculation loop. Plates were incubated at 26 °C and interactions recorded 5 d.p.i of *T. hamatum*.

2.6 Genome mining for secondary metabolite gene clusters

Sequencing of the *T. hamatum* GD12 genome was previously carried out in-house using Illumina sequencing technology. The genome was subsequently analysed for secondary metabolite gene clusters using the genome mining antiSMASH 2.0 platform[45]. A BLASTP analysis was performed on the output against the *Aspergillus nidulans* FGSC_A4 proteome (NCBI Taxonomy ID: 227321). The protein output was then aligned to the *A.nidulans* FGSC_A4 genome using the Integrative Genomics Viewer (IGV)[46] to determine the chromosomal location of secondary metabolite gene clusters.

2.7 LC-MS analysis

2.7.1 Sample preparation

Cultures of *S. sclerotiorum* M448 only, Δ*ThhepA*::*hph2* only, *T. hamatum* GD12 confronted with *S. sclerotiorum* M448 and Δ*ThhepA*::*hph2* confronted with *S. sclerotiorum* M448 were grown on PDA plates. After 48 h, the uncolonised media between the confrontation cultures, or the media surrounding the leading edge of lone cultures was extracted, along with non-inoculated PDA, made up to 50 mL with sterile milliQ water and incubated at 4 °C for 72 h. Extracts were then filtered through sterile miracloth and the supernatants collected, flash-frozen in liquid nitrogen and freeze-dried. Samples were re-suspended to 10 % of the original volumes (in this case 500 µL) in 10 % Methanol + 7.2 µg ml^{-1} umbeliferone as an internal standard, centrifuged for 10 min at 16,000 rpm, 4 °C, and filtered through a 0.2 µm filter on ice.

2.7.2 Analysis of secondary metabolite samples

Samples were run twice on LC-MS using a Polaris reversed phase C18 column – once in positive and once in negative ion modes, on an Agilent Quantitative-Time-Of-Flight (QTOF) mass spectrometer using electrospray ionisation. The LC-MS was run in full scan mode with tandem MS capabilities. Features were subsequently extracted using the Molecular Feature Extraction algorithm in Agilent's MassHunter software (Agilent Technologies, Germany) and the deconvoluted data aligned using an in-house Kernel Feature Alignment algorithm. Only features present in two out of three of the replicates were considered 'true'.

3. Results

3.1 Bioinformatics analysis of the *Trichoderma hamatum* GD12 genome

It has been hypothesised that regulation of gene clusters is dependent on chromosomal location[30], and gene clusters that regulate secondary metabolites are typically found within the subtelomeric region[9,30,47]. The antiSMASH output of the *T. hamatum* GD12 genome (Appendices 6.4) was subjected to BLASTP analysis against the extensively annotated genome of the closely related fungus *A. nidulans* FGSC_A4, and the protein sequence of the latter aligned against the genome in IGV. Figure 7 shows the distribution of SM gene clusters within the genome, with the height of the peak directly correlating to the number of secondary metabolite genes located within that region. It is clear that the majority of SM genes are located towards the ends of chromosomes, particularly chromosomes two and seven. As to whether or not these genes lay within the subtelomeric region would require further investigation into what depicts such a region.

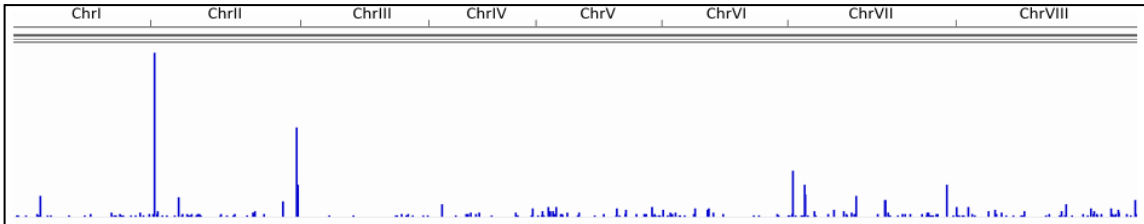


Figure 7 | Alignment of secondary metabolite gene clusters against the *Aspergillus nidulans* genome. The genome of *T. hamatum* GD12 was interrogated for secondary metabolite gene clusters using the genome mining software antiSMASH. BLASTP analysis was performed with the resulting protein output against the *A. nidulans* FGSC_A4 genome. The protein sequences from *A. nidulans* were aligned against its genome to determine the location of SM gene clusters. The height of each peak is relative to the number of genes found within that region. It is apparent that the majority of SM genes are located towards the end of chromosomes, particularly chromosomes two and seven.

3.2 Heterochromatin Protein 1 is highly conserved

HEP1 is a highly conserved protein with orthologs found in a broad range of eukaryotic organisms. Functionally, HEP1 is involved in chromatin re-modelling for activation and silencing of gene clusters. A HEP1 homolog was identified within the *T. hamatum* GD12 genome (Figure 8A). The coding sequence (CDS) for the HEP1 protein consists of 5 exons which are interrupted by 4 introns. HEP1 is defined by two highly conserved domains, an amino-terminal chromo domain and a carboxy-terminal chromo-shadow domain which are separated by a more variable linker region, thought to act as a hinge. The amino acid sequence of the GD12 homolog was aligned against an amino acid sequence from the closely related *T. atroviride* along with chromo domain and chromo-shadow domain sequences from *N. crassa, H. sapiens* and *D. melanogaster* using Clustal omega (Figure 8B). High levels of conservation are seen within the chromo domain (dark blue box) and chromo-shadow domain (light blue box). It was demonstrated in *Drosophila* that three aromatic residues (Tyr-24, Trp-45 and Tyr-48) are required to form a 'cage-like' structure creating a hydrophobic pocket to which methylation residues on H3K9 can bind[48]. The latter two of these aromatic residues were identified in GD12 (highlighted in red), however, the former appears to have either been misaligned (highlighted in aqua) or a possible SNP has taken place (indicated in green).

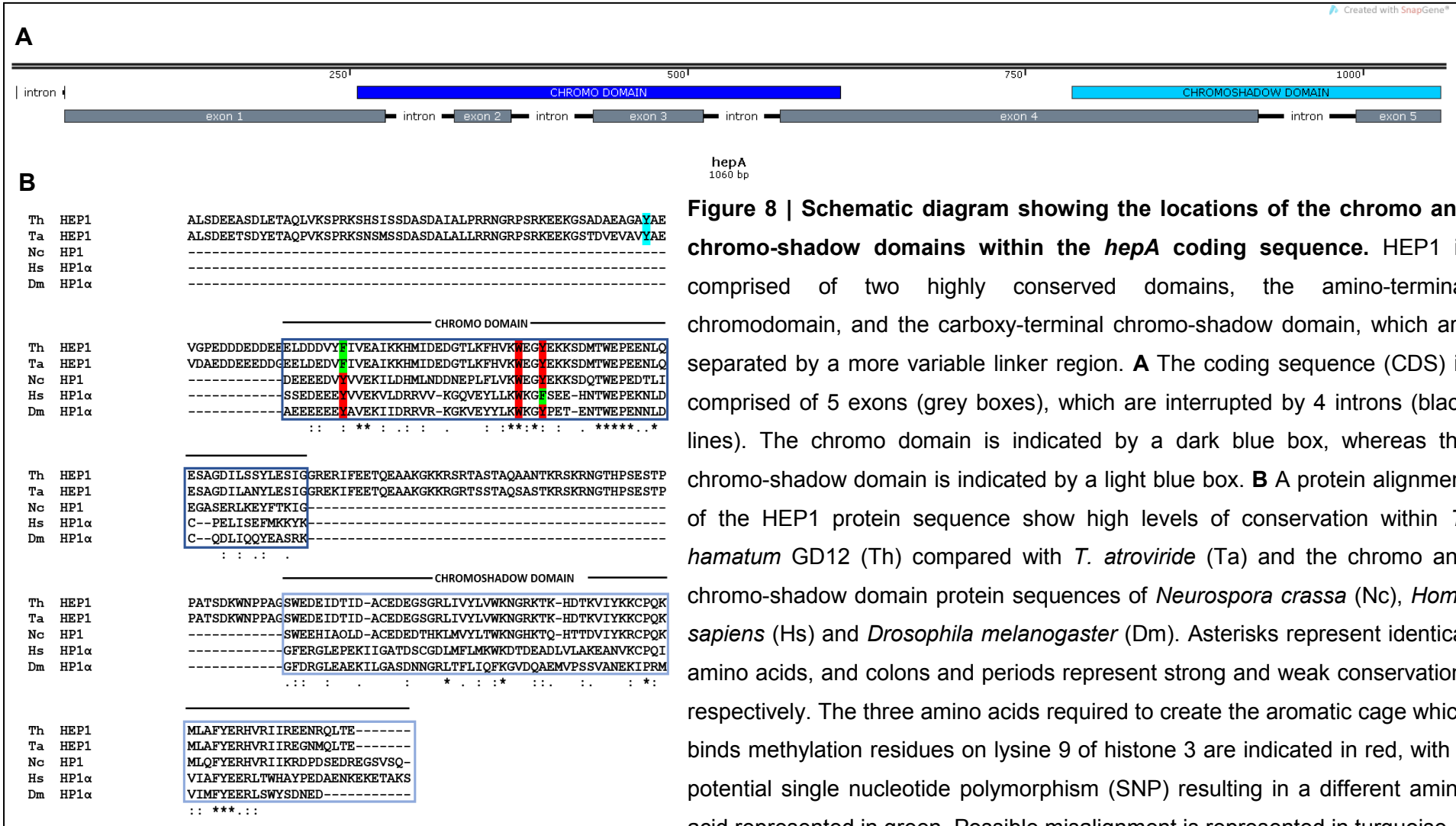


Figure 8 | Schematic diagram showing the locations of the chromo and chromo-shadow domains within the *hepA* coding sequence. HEP1 is comprised of two highly conserved domains, the amino-terminal chromodomain, and the carboxy-terminal chromo-shadow domain, which are separated by a more variable linker region. **A** The coding sequence (CDS) is comprised of 5 exons (grey boxes), which are interrupted by 4 introns (black lines). The chromo domain is indicated by a dark blue box, whereas the chromo-shadow domain is indicated by a light blue box. **B** A protein alignment of the HEP1 protein sequence show high levels of conservation within *T. hamatum* GD12 (Th) compared with *T. atroviride* (Ta) and the chromo and chromo-shadow domain protein sequences of *Neurospora crassa* (Nc), *Homo sapiens* (Hs) and *Drosophila melanogaster* (Dm). Asterisks represent identical amino acids, and colons and periods represent strong and weak conservation, respectively. The three amino acids required to create the aromatic cage which binds methylation residues on lysine 9 of histone 3 are indicated in red, with a potential single nucleotide polymorphism (SNP) resulting in a different amino acid represented in green. Possible misalignment is represented in turquoise.

3.3 Confirmation of Δ*ThhepA*::*hph* strains

Loss of HEP1 was achieved by using the split marker method of homologous recombination to replace the *hepA* ORF with the *hph* gene conferring hygromycin resistance (Figure 9).

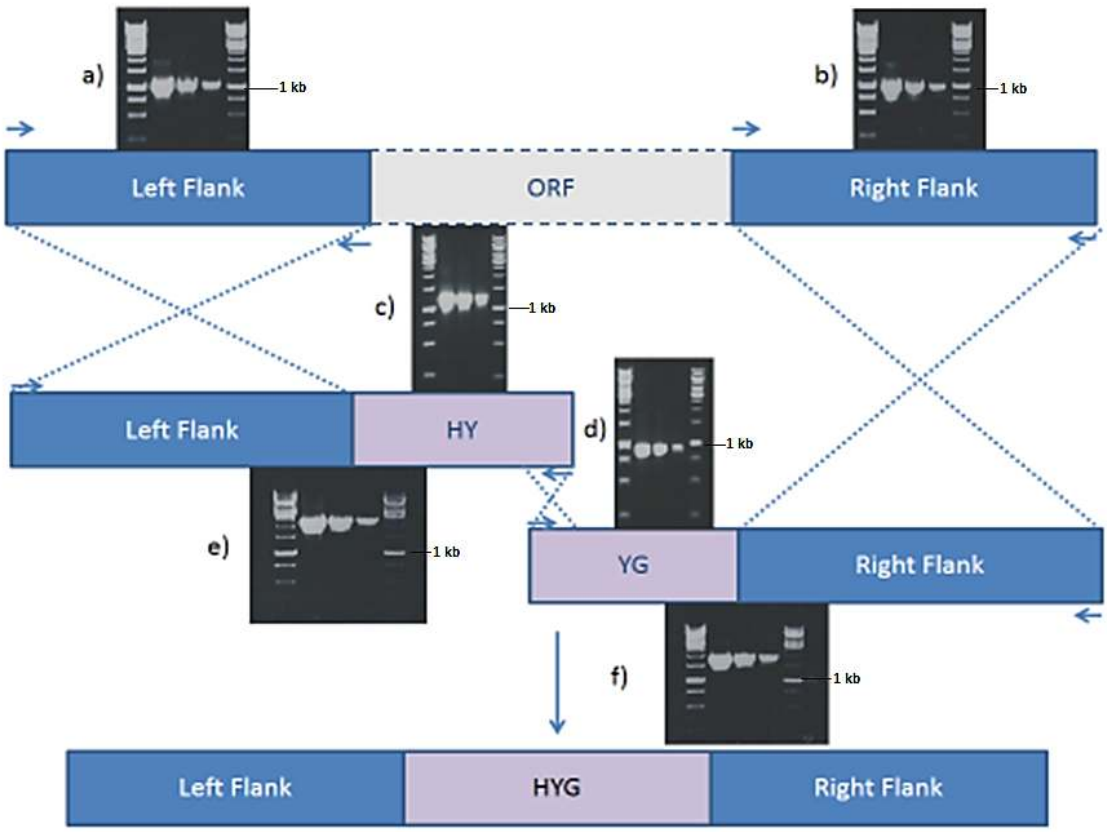


Figure 9 | Schematic diagram showing the split-marker method for creating the knockout cassette used for generating Δ*ThhepA*::*hph* strains. PCR amplicons were separated on 1.2 % agarose gels and were flanked by 10 µL and 5 µL of Fermentas 1 Kb GeneRuler ladder, respectively, unless stated otherwise. For quantification purposes, 10 µL, 5 µL and 2 µL of gel purified PCR products were loaded on to each gel respectively: a) left flanking region of *hepA* gene (1016 bp); b) right flanking region of *hepA* gene (940 bp); c) amplification of the former part of the *hph* gene (HY – 1200 bp); d) amplification of latter part of the *hph* gene (YG – 800 bp); e) second round Phusion® PCR product of LF + HY (2210 bp); f) second round Phusion® PCR product of RF + YG (1740 bp).

Confirmation of Δ*ThhepA*::*hph* strains was carried out by digestion of genomic DNA using the restriction enzyme *Stu*I (5'-AGG^CCT-3') (Appendices 6.3) and subsequent Southern blot analysis using a digoxigenin – dUTP (DIG) labelled probe of the left flanking region, ~1 Kb upstream, of the *hepA* open reading frame. A band size of 3201 bp confers a wild-type strain, whereas a band size of 6811 bp confers a *hepA* knockout mutant strain. The resulting Southern blot (Figure 10) confirms all putative *hepA* deletion mutants when compared to the wild-type strain GD12. Due to lack of another selectable marker, creation of a complementation strain is currently not possible, therefore, three independent *hepA* deletion mutants were selected for further experimentation and hereafter are referred to as Δ*ThhepA*::*hph*1, Δ*ThhepA*::*hph*2 and Δ*ThhepA*::*hph*3.

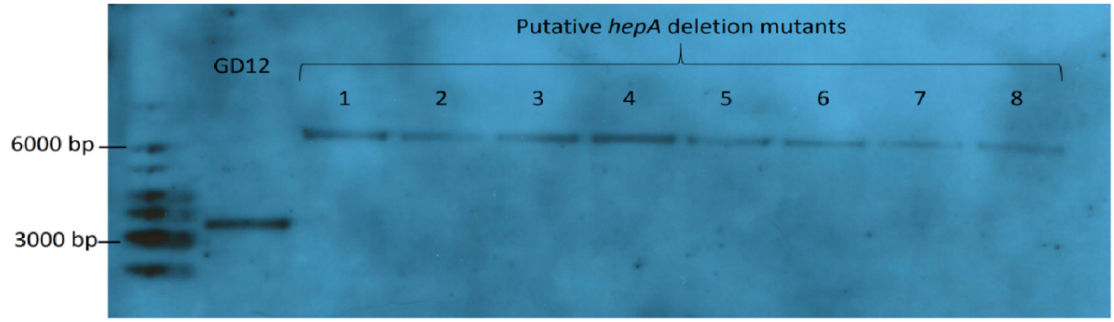


Figure 10 | Confirmation of *hepA* deletion via DIG-Southern blot analysis. Genomic DNA was extracted from putative *hepA* deletion mutants and subsequently digested using restriction enzyme *Stu*I. A DIG-labelled left flank probe was created to probe for wild-type strains (3201 bp) and *hepA* deletion strains (6811 bp). Lane 1; Fermentas GeneRuler 1 Kb ladder, lane 2; GD12 genomic DNA, lanes 3-10; putative *hepA* deletion mutants.

3.4 Phenotypic analysis of Δ*ThhepA*::*hph* strains

3.4.1 Loss of HEP1 leads to growth inhibition

Phenotypic analysis of the Δ*ThhepA*::*hph* strains was carried out on PDA plates under both light and dark conditions. Hyphal growth was recorded over a 10-day period by scanning each plate with an Epson Perfection V750 Pro scanner and images analysed using imageJ. The three independent Δ*ThhepA*::*hph* strains all showed significantly compromised hyphal growth compared with the wild-type strain GD12 (Figure 11-1). Growth of the mutants was sporadic, with hyphal proliferation below the agar surface. The wild-type strain GD12, however, grew in a more consistent manner, with an even distribution, reaching the edge of the agar plate within 4 – 5 days (Figure 11-2). These findings are consistent with a spore count carried out on PDA plates grown in light and dark conditions for 14 days (Figure 11-3). The Δ*ThhepA*::*hph* mutants had a significantly reduced spore count, compared with the wild-type, under both light (Figure 11-3A) and dark (Figure 11-3B) growth conditions.

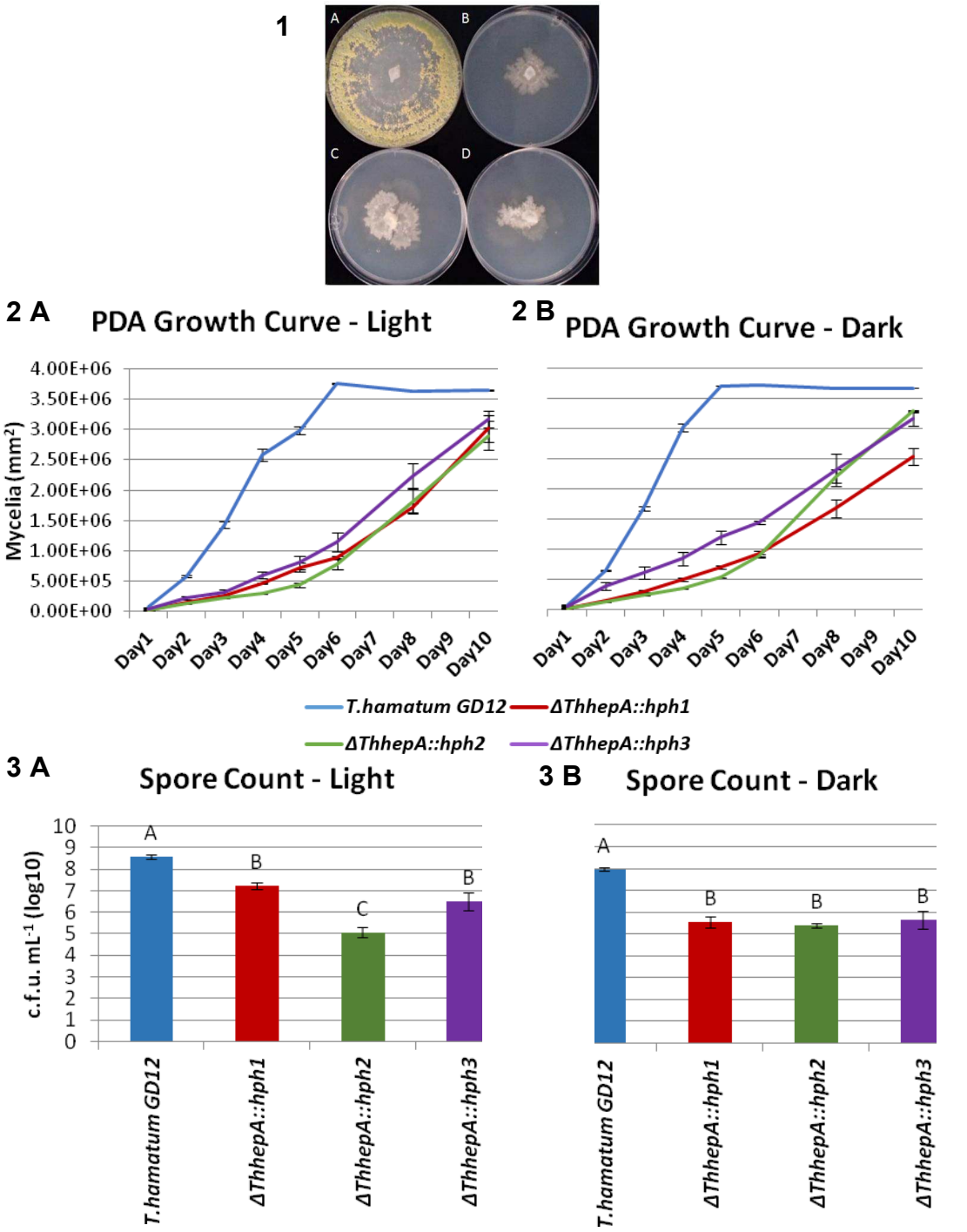


Figure 11 | Phenotypic analysis of *hepA* mutants. Loss of HEP1 leads to growth inhibition of mutant strains (**1**: **A** – wild-type GD12, **B** - Δ*ThhepA*::*hph*1, **C** - Δ*ThhepA*::*hph*2, **D** - Δ*ThhepA*::*hph*3. Analysis of growth rates on PDA plates in light (**2A**) and dark (**2B**) shows mutants have significantly reduced growth compared to GD12 over a 10 day growth period. Bars are the means of replicates ± standard errors. This is consistent with spore counts at 14 d.p.i. (**3A** – light, **3B** – dark) with *hepA* mutants producing significantly fewer spores. Bars are the means of replicates ± standard errors. Two-way ANOVA and post-hoc Tukey tests show significant differences ($p<0.05$) indicated by different letters.

3.4.2 Loss of HEP1 leads to changes in antimicrobial activity

To investigate antimicrobial activities concomitant with loss of HEP1, confrontation plates were established by inoculating the upper half of a PDA plate with *T. hamatum* and the lower half with a range of both plant and human pathogens. When *T. hamatum* GD12 was confronted with a range of plant pathogenic fungi and oomycetes, the wild-type strain rapidly overgrew the pathogens. (Figure 12). In contrast, Δ*ThhepA*::*hph* mutants displayed antibiosis producing a zone of inhibition surrounding mutant colonies.

Confrontation assays were also performed with a range of multidrug resistant human pathogenic yeasts. The wild-type strain GD12 similarly overgrew the yeasts, whereas Δ*ThhepA*::*hph* strains displayed a zone of growth inhibition (Figure13).

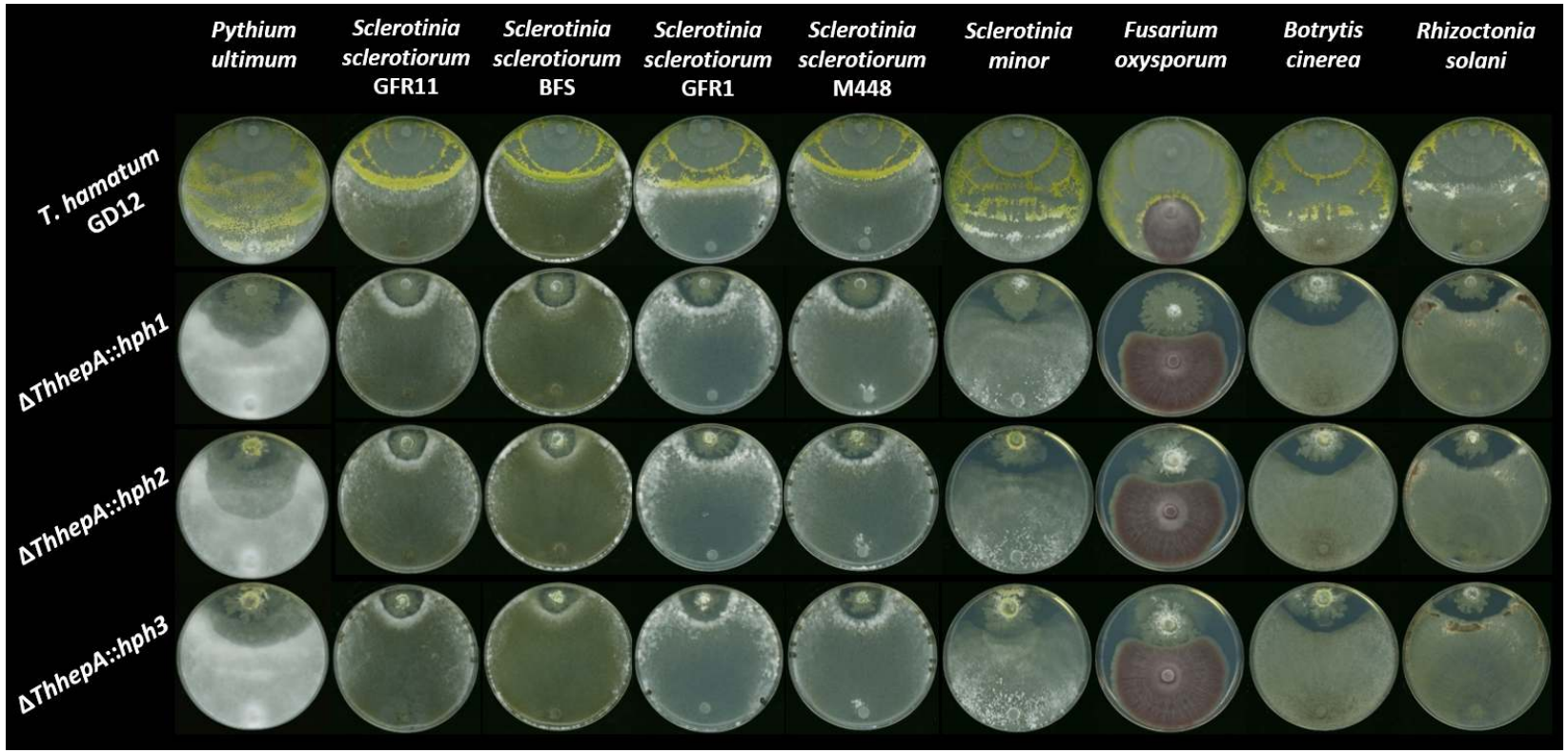


Figure 12 | Broad-spectrum inhibition of pathogenic fungi and oomycetes. The upper half of PDA plates were inoculated with wild-type *T. hamatum* GD12 (top row) and three independent *hepA* deletion mutants Δ*ThhepA*::*hph*1, Δ*ThhepA*::*hph*2 and Δ*ThhepA*::*hph*3 (bottom three rows respectively). The lower half of the PDA plates were inoculated with various plant pathogenic fungi. After 5 days growth, the interactions between *T. hamatum* strains and pathogens were recorded. *T. hamatum* GD12 displays inhibition of the pathogens by hyphal overgrowth. Loss of HEP1, however, results in a zone of inhibition (antibiosis) of the pathogens.

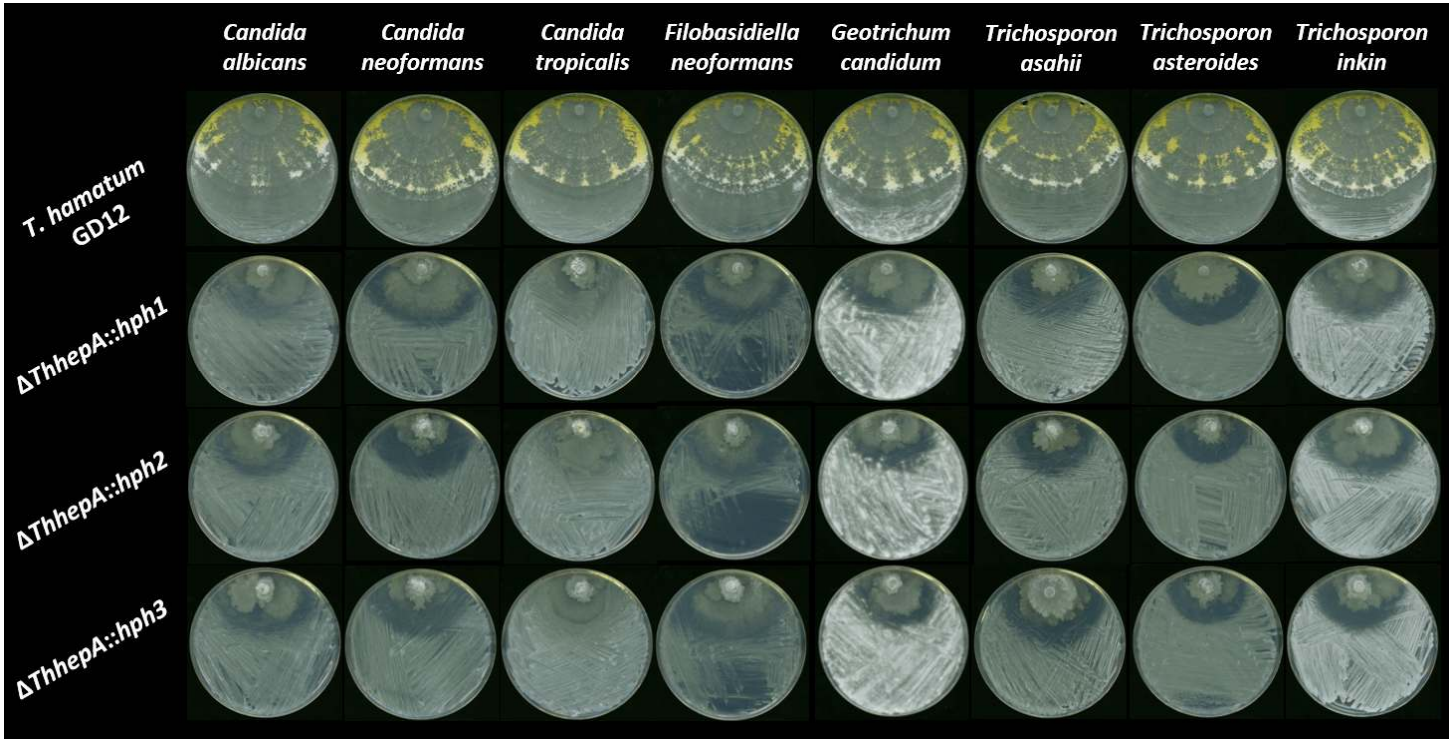


Figure 13 | Broad spectrum inhibition of multi-drug resistant human pathogenic yeasts. The upper half of PDA plates were inoculated with wild-type *T. hamatum* GD12 (top row) and three independent *hepA* deletion mutants Δ*ThhepA*::*hph*1, Δ*ThhepA*::*hph*2 and Δ*ThhepA*::*hph*3 (bottom three rows respectively). *T. hamatum* strains were allowed to grow for 3 days before any un-colonised area of the plates were streaked with a range of human pathogenic yeasts. Strains were allowed to grow for a further 2 days before interactions were recorded. *T. hamatum* GD12 displays inhibition of the pathogens by hyphal overgrowth .Loss of HEP1, however, results in a zone of inhibition (antibiosis) of the pathogens.

3.4.3 Loss of HEP1 leads to an altered secretome

To determine if loss of HEP1 leads to an altered metabolome, samples were taken from the media surrounding of the leading edge of actively growing fungi. Samples consisted of; Δ*ThhepA*::*hph*2 and *S. sclerotiorum* M448 independent strains, as well as the uncolonised media from confrontation and inhibition zones of *T. hamatum* GD12 and Δ*ThhepA*::*hph*2, respectively, when confronted with *S. scleroriorum* M448 (2 d.p.i). A non-inoculated PDA plate was also analysed and any features present in the sample were considered background and subsequently removed from all other sample sets.

When data is visualised as a heat map, feature clustering is consistent between replicates for each treatment type, both in positive ionisation mode (Figure 14 A) and negative ionisation mode (Figure 14 B). Multiple features are up- and down- regulated differentially between treatment types, yet others are constitutively expressed.

Visualising the data through Venn diagrams (Figure 15) shows clear clustering of features. The number of 'core' features in both data sets is very low, with 26 shown in positive ionisation mode and only 4 with negative ionisation. Loss of HEP1 leads to constitutive expression of 45 positive ionisation features and 30 negative. Interestingly, 301 positive and 125 negative ionisation features are unique to the Δ*ThhepA*::*hph*2 sample, whereas a further 95 positive and 116 negative ionisation features are only expressed when Δ*ThhepA*::*hph2* is confronted with *S. sclerotiorum* M448. Deletion of *hepA* leads to the loss of 82 positive and 100 negative ionisation features which are only present during *T. hamatum* GD12 confrontation with *S. sclerotiorum* M448. Ionisation features which are considered 'core' to the *T. hamatum* secretome indicate 23 positive

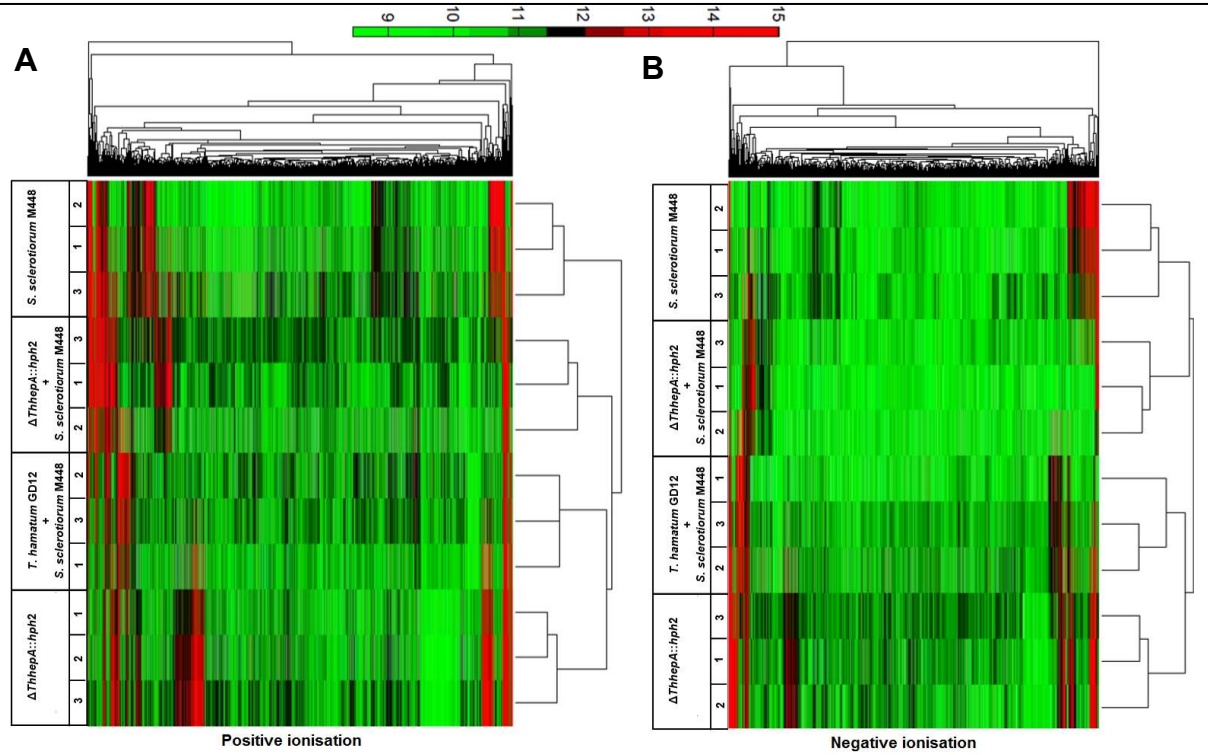


Figure 14 | Heat maps showing differential fingerprint clustering of secreted compounds produced by *T. hamatum* and *S. sclerotiorum* M448. Samples containing secreted compounds were taken from the inhibition/interaction zone of confrontation plates, or from media surrounding the leading edge of actively growing independent cultures and were run twice on an LC-MS with a polaris C18 reversed phase column (once in positive ion mode (**A**) and once in negative (**B**)) on QTOF using electrospray ionisation. Features which were found in two out of three of the replicates were considered genuine. Clustering within each of the treatment types is consistent and differential expression is seen between the individual treatment types.

and 11 negative features, which are present across all *T. hamatum* strains whether they are confronted with *S. sclerotiorum* or not. However, features present only when *T. hamatum* strains are confronted with *S. sclerotiorum* reveals 17 positive and 15 negative. From here we began an attempt to identify some key features secreted by Δ*ThhepA*::*hph*2 which may display antimicrobial activities. Due to two of the sample vials becoming damaged whilst being analysed on the LC-MS, replicates 2 of GD12-M448 and Δ*ThhepA*::*hph*-M448 were removed.

Key features were normalised to the internal standard, umbelliferone, and relative abundance presented as bar charts. Positive ionisation data (Figure 16) revealed an interesting array of key features, many of which had characteristic pharmaceutical potential. **Dibenzo-quinoline carbaldehyde** (Figure 16 A) was highly expressed in PDA with lower levels of expression present in all other samples, particularly GD12-M448. Levels of **tetrahydro-quinoline carbaldehyde** (Figure 16 B) were abundant in samples containing *hepA* mutants, but did not appear to be present, at least not in significant abundance, in all other samples. **Brefeldin A** (Figure 16 C) and **adephenine** (Figure 16 F) were in low abundance in *S. sclerotiorum* M448 only samples, in high abundance when in confrontation with *hepA* mutants, but absent when in confrontation with the wild-type GD12. **Norcantharidin** (Figure 16 D), **kynurenic acid** (Figure 16 E) and **1-methyl-2-pyrrolidone** (Figure 16 G) were only present in samples containing *hepA* mutants, and were particularly abundant when in confrontation with *S. sclerotiorum* M448. Low levels of **desthiobiotin** (Figure 16 I) were recorded when GD12 was confronted with *S. sclerotiorum* M448 and when *hepA* mutants were grown independently, with slightly higher levels being presented in all other samples. Constitutively low

levels of **(s)-(-)-perillic acid** expression were displayed across all sample sets, with the exception of *hepA* mutants confronted with *S. sclerotiorum* M448 where elevated levels were shown.

Analysis of key negative ion features showed constitutively high levels of **4-hydroxy-3-methoxybenzoate** expression (Figures 17 B (3.8 min RT), C (4.3 min RT) and D (8.8 min RT)) in all samples containing *hepA* mutants, but was absent from all other sample sets. Levels of **camptothecin** (Figure 17 A) and **sebacate** (Figure 17 G) were low across all samples, with slightly elevated expression in samples containing *hepA* mutants, particularly in *hepA* mutants as independent cultures. Confrontation samples of Δ*ThhepA*::*hph* with *S. sclerotiorum* M448 displayed elevated abundance levels of **N-methylanthranilate** (Figure 17 E), with slightly lower levels displayed in Δ*ThhepA*::*hph* only samples. However, it was absent from all other treatments. **Phenylalanine** (Figure 17 H) was relatively abundant across all sample sets, with lower levels expressed in *S. sclerotiorum* M448 only samples, and when in confrontation with GD12. Independent Δ*ThhepA*::*hph* strains displayed the lowest levels of phenylalanine. Low levels of **2-deoxyribose 5-phosphate** (Figure 17 F) were recorded in PDA, *S. sclerotiorum* M448 and Δ*ThhepA*::*hph* samples, with slightly higher levels found in confrontation samples.

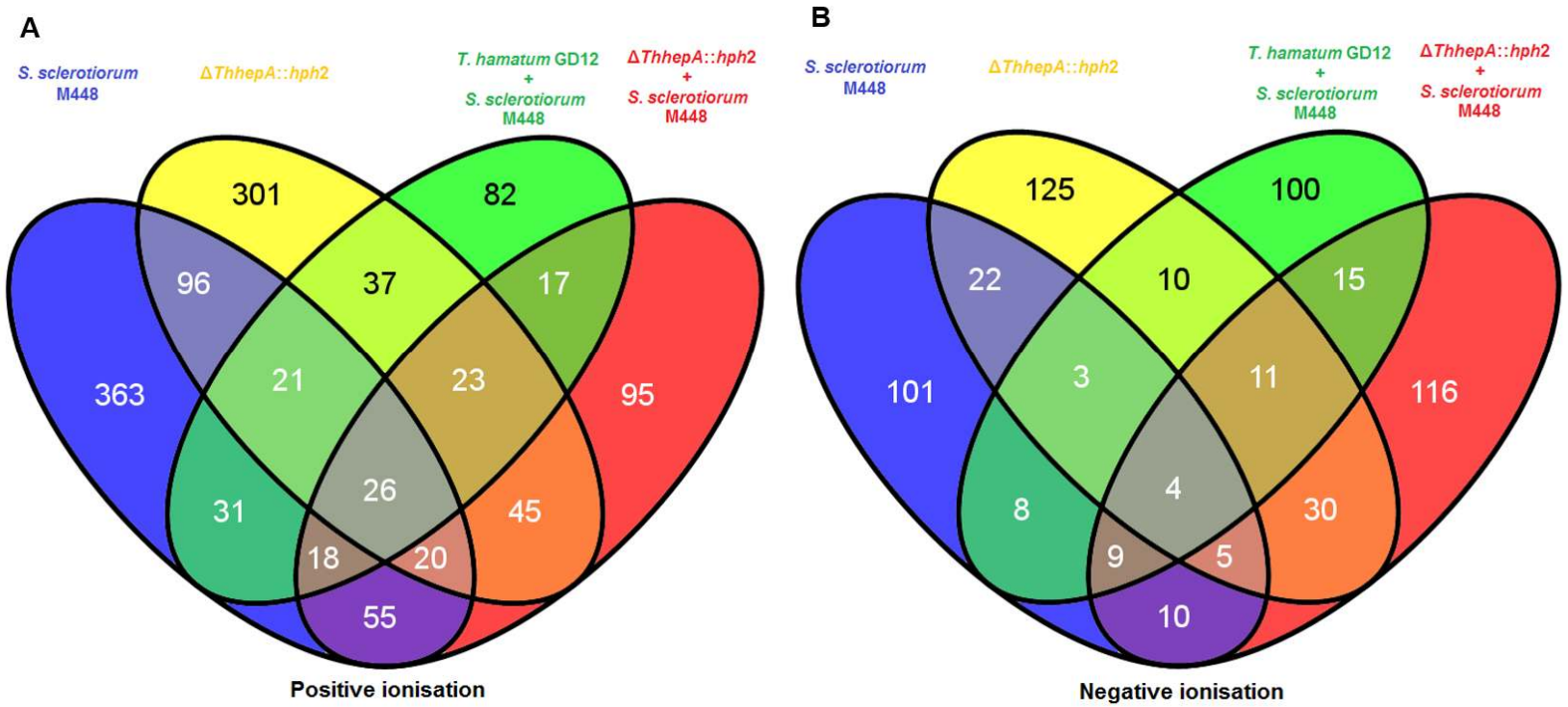


Figure 15 | Venn diagrams showing clustering of secreted features identified by using LC-MS (QTOF) analysis. Data from positive ionisation (**A**) and negative ionisation (**B**) heat maps (Figure 14) was converted into Venn diagrams to visualise differentiation of secreted compounds. Blue represents *S. sclerotinia* M448 only, yellow represents Δ*ThhepA*::*hph*2 only, green represents *T. hamatum* GD12 confronted with *S. sclerotiorum* M448 and red represents Δ*ThhepA*::*hph*2 confronted with *S. sclerotiorum*. As all strains were grown on PDA, any features found in the PDA only sample were considered background and were subsequently removed from all other data sets.

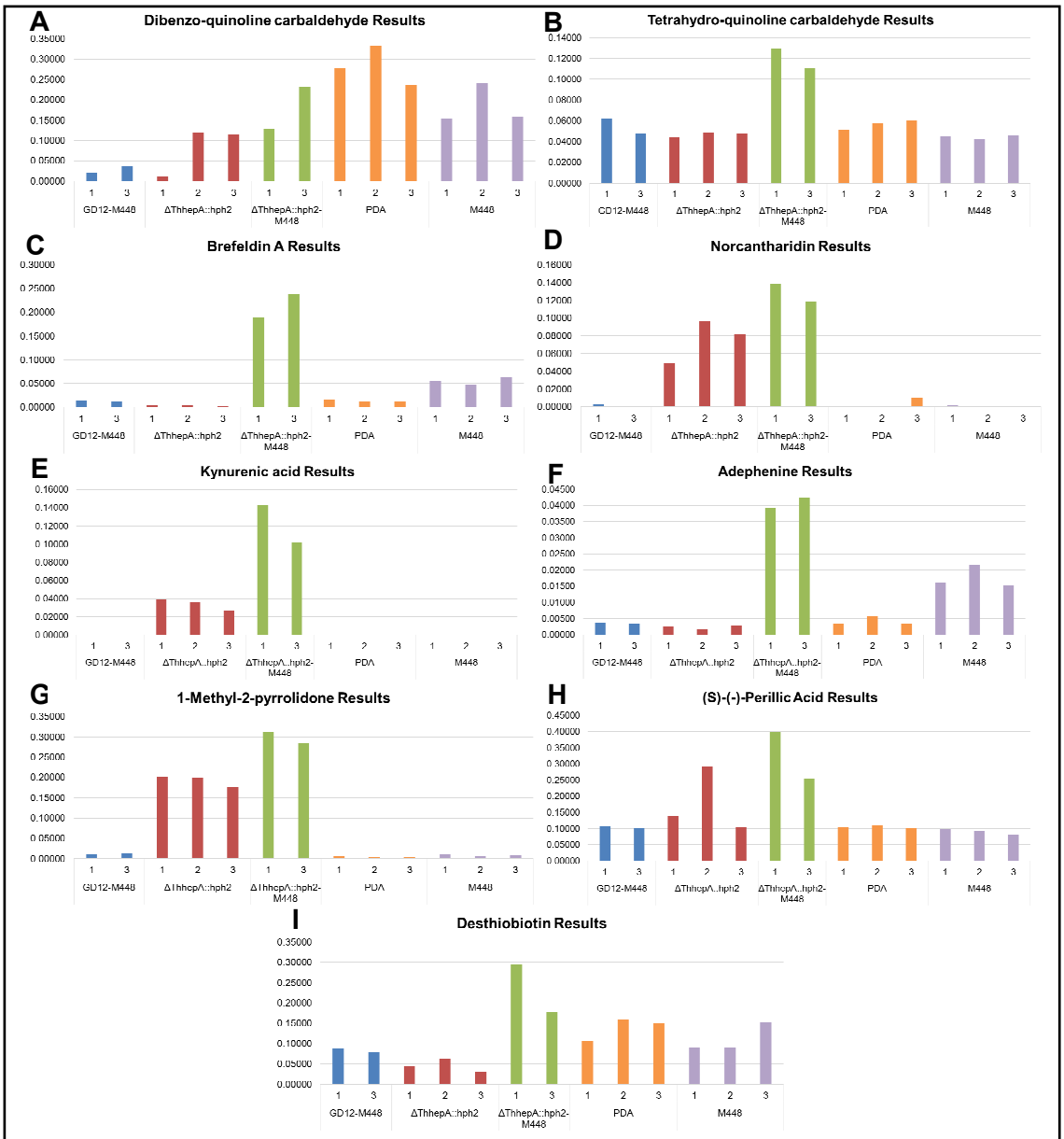


Figure 16 | Key features identified from positive ionisation samples. Treatment types are as follows: GD12-M448 (blue), Δ*ThhepA*::*hph* (red), Δ*ThhepA*::*hph*-M448 (green), PDA (orange) and M448 (purple). (**A**) Dibenzo-quinoline carbaldehyde; (**B**) Tetrahydro-quinoline carbaldehyde – used in the production of photosensitive materials; (**C**) Brefeldin A – an antimicrobial agent; (**D**) Norcantharidin – conveys anticancer properties; (**E**) Kynurenic acid – an anticonvulsant; (**F**) Adephenine – a smooth muscle relaxant; (**G**) 1-Methyl-2-pyrrolidone – used in the recovery of pure hydrocarbons; (**H**) (S)-(-)-Perillic acid – a hypoglycemic agent used as an anti-diabetic; (**I**) Desthiobiotin – an immunosuppressive agent.

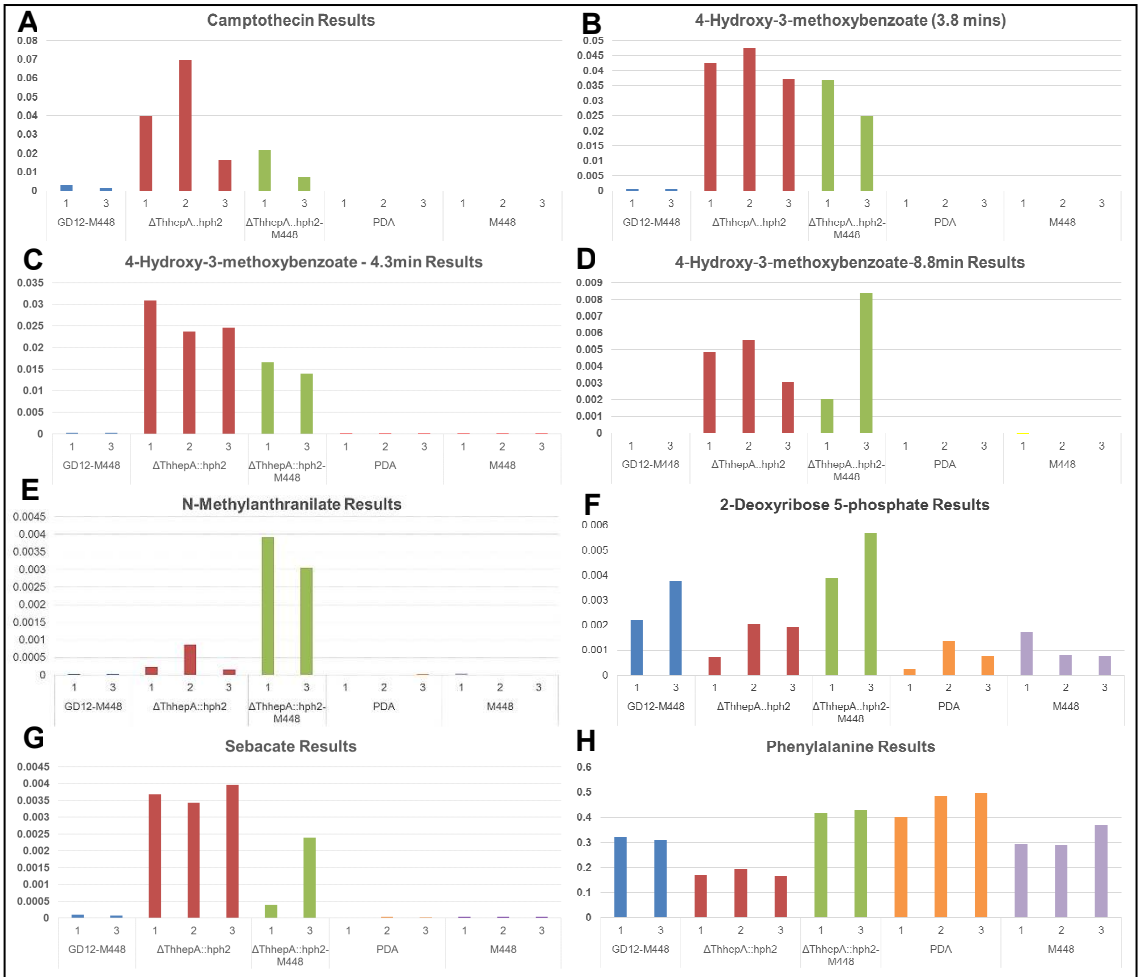


Figure 17 | Key features identified from negative ionisation samples. Treatment types are as follows: GD12-M448 (blue), Δ*ThhepA*::*hph* (red), Δ*ThhepA*::*hph*-M448 (green), PDA (orange) and M448 (purple). (**A**) Campothecin – inhibits topoisomerase I; (**B**, **C** and **D**) 4-Hydroxy-3-methoxybenzoate – a precursor for Vanillin, a vanilla bean extract; (**E**) N-Methylanthranilate – the starter unit of acridone alkaloid biosynthesis, bioactive property as an antimalarial drug; (**F**) 2-Deoxyribose 5-phosphate – catalyzes the reverse aldol condensation reaction without any cofactors; (**G**) Sebacate – used in the manufacturing of polymer plastics; (**H**) Phenylalanine – function unclear, potential use as a pharmacological agent.

4. Discussion

As the human population increases and alternative approaches to sustainable agricultural intensification are sought, identifying new sources of bioactive compounds that promote plant productivity and inhibit plant and humans pathogenic fungi is of major importance for both food security and for human health.

This project set out to determine whether deletion of HEP1 in the plant-growth-promoting and biocontrol fungus *T. hamatum* GD12 might lead to activation of cryptic secondary metabolite gene clusters and concomitant secretion of novel bioactive compounds with antimicrobial activities. Indeed, results show that loss of the HEP1 protein leads to altered antimicrobial activity with simultaneous changes in the fungal secretome. These findings directly correlate to the differential fingerprint clustering derived from LC-MS analysis.

Bioinformatics analysis of HEP1 indicates high levels of conservation of GD12 with *T. atroviride* (87.64% homology). The chromo and chromo-shadow domains are particularly highly conserved when compared with other eukaryotic organisms, suggesting the function of HEP1 in *T. hamatum* GD12 is consistent with that of other eukaryotic organisms. Binding of the chromo domain to di- and tri- methylation residues on H3K9 is achieved by three conserved amino acids, Tyr-24, Trp-45 and Tyr-48 in *Drosophila*, creating a hydrophobic cage acting as the binding site. The latter two of these amino acids are conserved in GD12, however, the former appears to be either misaligned, or a SNP has modified the UAC codon which translates to Tyr into UUC which conferring Phe. Further investigation would be required to determine the exact structure of the *T. hamatum* homolog. Nevertheless, investigations conducted here suggest a

function of the protein in chromatin re-modelling and regulation of secondary metabolite gene clusters.

Altered antimicrobial activity as a consequence of altered gene cluster regulation was established through confrontation assays against a range of both agriculturally and medically important pathogens (displaying multiple host range and multidrug resistance, respectively), with loss of HEP1 leading to aberrant growth of mutants, and antibiosis of pathogens. Analysis of the small molecule secretome via LC-MS analysis suggests HEP1 may be a 'hub' gene, with involvement in multiple pathways. Many features from both positive and negative ionisation data sets are differentially up- and down- regulated due to loss of HEP1. Interestingly, however, 82 positive and 100 negative features are present when *T. hamatum* GD12 is confronted with *S. sclerotiorum* M448 which are not present in any other samples, suggesting that HEP1 is indeed involved in multiple pathways and either the protein itself or bioactive product(s) derived from the pathway may subsequently repress other SM gene clusters. However, confirmation of this hypothesis would require further investigation.

Key features identified from both positive and negative ion LC-MS data reveal a broad range of differentially secreted proteins. Constitutively high levels of (S)-(-)-perillic acid (causing growth inhibition and protein prenylation of cancer cells[49]) and 4-hydroxy-3-methoxybenzoate (RT; 3.8 min, 4.3 min and 8.8 min – bioactive characteristics uncharacterised) are expressed in samples containing *hepA* mutants, with levels of expression decreasing as retention time increases for the latter. This may suggest HEP1 is a key regulator of the biosynthetic pathways from which these compounds are derived. Other data suggests HEP1 is a regulator of N-methylanthranilate (a bird repellent used on grasses[50] which is also found in grapes[51]), kynurenic acid (a product of L-tryptophan metabolism

with high levels found in patients suffering from tick-borne encephalitis[52] and schizophrenia[53]), 1-methyl-2-pyrrolidone (used in the recovery of hydrocarbons from petrochemical processing[54] but has recently been identified as a reproductive toxicant[55]) and norcantharidin (a demethylated analogue of the natural toxin – cantharidin, shown to inhibit growth of tumors[56]) are only present through deletion of *hepA*, but are further upregulated when Δ*ThhepA*::*hph* strains are confronted with *S. sclerotiorum* M448. These findings suggest HEP1 does indeed play a role in regulation of gene clusters. However, some identified compounds are further up-regulated during confrontation with *S. sclerotiorum* M448, suggesting other regulatory proteins may also play a key role in the biosynthesis of product(s) derived from the pathway(s). Camptothecin is a topoisomerase I inhibitor[57] which has been utilised for anticancer treatment[58] and is found in abundance in Δ*ThhepA*::*hph* samples, with slight down-regulation when in confrontation with *S. sclerotiorum* M448. A similar pattern is seen with sebacate (used in the production of polymers for targeted drug delivery[59]).

All of the above compounds are unique to samples containing Δ*ThhepA*::*hph* strains, supporting the hypothesis that *hepA* may be a 'hub' gene with involvement in multiple pathways. Another identified feature of particular interest is Brefeldin A, which is highly abundant when Δ*ThhepA*::*hph* is confronted with *S. sclerotiorum* M448, but absent from all other sample sets. This lactone antibiotic inhibits GBF1 (guanine nucleotide exchange factor) causing movement of secretory proteins from the golgi into the endoplasmic reticulum (ER) which in turn activates ER stress and results in apoptosis[60,61]. When considering the mutant-pathogen confrontation assays, a clear zone of inhibition

is displayed surrounding the leading edge of Δ*ThhepA*::*hph* cultures. Secretion of Brefeldin A could provide an explanation for this phenotype.

Due to increased demand for novel bioactive compounds[62–64], successful identification and application would prove beneficial. Compounds effective as fungicides against plant diseases should ideally provide systemic resistance. Brefeldin A was shown to be highly upregulated in Δ*ThhepA*::*hph* + *S. sclerotiorum* M448 LC-MS samples inhibits membrane transport and is therefore toxic to eukaryotic cells and would not be a useful candidate[60,65].

Over use of antibiotics has led to an increase in drug-resistant fungal pathogens and the emergence of fungal pathogens with intrinsic resistance to mould-active compounds, means that novel antifungal compounds are urgently needed to control opportunistic fungal infections in the ever-increasing population of immunocompromised patients[63]. With no new antifungal drugs on the immediate horizon and with azole resistance now becoming widespread in hospitals, identification of novel antimicrobial compounds would therefore be of enormous benefit to medicine by providing alternatives to the azoles, echinocandins and polyenes that inhibit fungal cell wall biosynthesis and which display varying activities against fungi capable of causing human infections. Further investigations of the regulatory pathways governing secondary metabolite biosynthesis in naturally occurring soil fungi such as *Trichoderma* spp. may allow the discovery of previously uncharacterised antimicrobial compounds with new modes of action

4.1 Concluding remarks and future investigation

Loss of HEP1 has a significant impact on the growth and development of *T. hamatum* GD12, an alteration in the biocontrol capabilities of the fungus, and most significantly an alteration to the secretome.

This project has begun to investigate bioactive product(s) derived from the biosynthetic pathway in which HEP1 is a regulatory component. The research has also indicated the possible involvement of the HEP1 protein in additional regulatory pathways. The project could have benefited from RT-PCR to determine the effects that the loss of HEP1 has on other genes involved in secondary metabolite biosynthesis. This might also identify other potential genes of interest that have a significant role in the production of secondary metabolites. Demonstrating antibiosis of pathogens using metabolite extracts, lacking the presence of actively growing *T. hamatum*, would confirm the phenotype displayed is caused by stable secretion compounds, and that the presence of *T. hamatum* is not required for its function.

All confrontation assays conducted within this study were carried out on PDA, however, this alone may have a significant impact on the efficacy of *Trichoderma* to display biocontrol properties. Therefore, investigating this phenotype on a range of minimal and rich media would allow a better understanding.

Furthermore, the zone of inhibition phenotype portrayed by the mutant strains may include an element of autolysis. This may be investigated through microscopy and/or release of a cytoplasmic marker enzyme.

Creation of a complemented strain of Δ*ThhepA*::*hph*2 would confirm that the phenotypes displayed by loss of HEP1 are due to loss-of-function, by restoring the phenotype displayed by the wild-type through re-insertion of *hepA* into the genome. Although previous attempts have proved this to be difficult due to lack of an additional selectable marker, further study would clearly benefit from such a strain and investigation into other transformation techniques may yield more success in this area.

Analysis of the secretome by LC-MS has begun to reveal some interesting results. Testing the biocontrol activities of the identified key features would significantly benefit this study. Also, analysis of a *T. hamatum* GD12 only sample set would allow confirmation of constitutive expression from Δ*ThhepA*::*hph*2 which is not present in the wild-type. However, the focus of this project was to identify novel antimicrobial compounds and therefore, a GD12 only sample set was not investigated. All of the above mentioned compounds were isolated and identified from a methanol extraction method, however, this may limit detection of more hydrophobic compounds. Further investigation may benefit from a chloroform extraction also, for further identification of non-polar compounds.

Overall, this project has demonstrated that genetic modification is a useful resource for the identification of novel compounds, and the data and resources generated from this study constitutes a strong foundation for further research into this subject area.

5. References

1. Reino, J. L., Guerrero, R. F., Hernández-Galán, R. & Collado, I. G. Secondary metabolites from species of the biocontrol agent *Trichoderma*. *Phytochem. Rev.* **7,** 89–123 (2007).

2. Bhetariya, P. J., Madan, T., Basir, S. F., Varma, A. & Usha, S. P. Allergens/Antigens, toxins and polyketides of important *Aspergillus* species. *Indian J. Clin. Biochem.* **26,** 104–19 (2011).

3. York, N. E. W., Division, U. N. P. & Nations, U. Embargoed until 12□: 00 PM , 11 March , 2009 WORLD POPULATION TO EXCEED 9 BILLION BY 2050□: (2009).

4. Carlet, J., Jarlier, V., Harbarth, S., Voss, A., Goossens, H. & Pittet, D. Ready for a world without antibiotics□? The Pensières Antibiotic Resistance Call to Action. *Antimicrob. Resist. Infect. Control* **1,** 11 (2012).

5. Clardy, J., Fischbach, M. & Currie, C. The natural history of antibiotics. *Curr. Biol.* **19,** 1–8 (2010).

6. Challis, G. L. & Hopwood, D. A. Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *Proc. Natl. Acad. Sci. U. S. A.* **100 Suppl ,** 14555–61 (2003).

7. Fischbach, M. A. Antibiotics From Microbes: Converging To Kill. *Curr. Opin. Microbiol.* **12,** 520–527 (2009).

8. Yu, J.-H. & Keller, N. Regulation of secondary metabolism in filamentous fungi. *Annu. Rev. Phytopathol.* **43,** 437–58 (2005).

9. McDonagh, A. *et al.* Sub-telomere directed gene expression during initiation of invasive aspergillosis. *PLoS Pathog.* **4,** e1000154 (2008).

10. Perrin, R. M., Fedorova, N. D., Bok, J. W., Cramer Jr, R. A., Wortman, J. R., Kim, H. S., Nierman, W. C. & Keller, N. P. Transcriptional regulation of chemical diversity in *Aspergillus fumigatus* by LaeA. *PLoS Pathog.* **3,** e50 (2007).

11. Windham M.T., Elad Y., B. R. A Mechanism for Increased Plant Growth Induced by *Trichoderma* spp. *Phytopathology* **76,** 518–521 (1986).

12. Ya-Chun Chang, Yih-Chang Chang, R. B. Increased Growth of Plants in the Presence of the Biological Control Agent *Trichoderma harzianum*. *Am. Phytopathol. Soc.* **70,** 145–148 (1986).

13. Brotman, Y., Lisec, J., Méret, M., Chet, I., Willmitzer, L. & Viterbo, A. Transcript and metabolite analysis of the *Trichoderma*-induced systemic

resistance response to *Pseudomonas syringae* in *Arabidopsis thaliana*. *Microbiology* **158,** 139–46 (2012).

14. Horst, L. E., Locke, J. & Krause, C. R. Suppression of *Botrytis* Blight of Begonia by *Trichoderma hamatum* 382 in Peat and Compost-Amended Potting Mixes. *Am. Phytopathol. Soc.* **89,** 1195–1200 (2005).

15. Segarra, G., Casanova, E., Avilés, M. & Trillas, I. *Trichoderma asperellum* strain T34 controls *Fusarium* wilt disease in tomato plants in soilless culture through competition for iron. *Microb. Ecol.* **59,** 141–9 (2010).

16. Geraldine, A. M., Lopes, F. A. C., Carvalho, D. D. C., Barbosa, E. T., Rodrigues, A. R., Brandão, R. S., Ulhoa, C. J. & Junior, M. L. Cell wall-degrading enzymes and parasitism of sclerotia are key factors on field biocontrol of white mold by *Trichoderma* spp. *Biol. Control* **67,** 308–316 (2013).

17. Vinale, F., Sivasithamparam, K., Ghisalberti, E. I., Marra, R., Barbetti, M. J., Li, H., Woo, S. L. & Lorito, M. A novel role for *Trichoderma* secondary metabolites in the interactions with plants. *Physiol. Mol. Plant Pathol.* **72,** 80–86 (2008).

18. Studholme, D. J., Harris, B., Le Cocq, K., Winsbury, R., Perera, V., Ryder, L., Ward, J. L., Beale, M. H., Thornton, C. R. & Grant, M. Investigating the beneficial traits of *Trichoderma hamatum* GD12 for sustainable agriculture — insights from genomics. *Front. Plant Sci.* **4,** 1–13 (2013).

19. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8,** 785–6 (2011).

20. Smith, D. J., Burnham, M. K. R., Bull, J. H., Hodgson, J. E., Ward, J. M., Browne, P., Brown, J., Barton, B., Earl, A. J. & Turner, G. B-Lactam antibiotic biosynthetic genes have been conserved in clusters in prokaryotes and eukaryotes. *EMBO J.* **9,** 741–747 (1990).

21. Lo, H-C., Entwistle, R., Guo, C-J., Ahuja, M., Szewczyk, E., Hung, J-H., Chiang, Y-M., Oakley, B. R. & C. C. C. Wang. Two separate gene clusters encode the biosynthetic pathway for the meroterpenoids, austinol and dehydroaustinol in *Aspergillus nidulans*. *J Am Chem Soc.* **134,** 4709–4720 (2012).

22. Zheng, X-Y., Spivey, N. W., Zeng, W., Liu, P-P., Fu, Z. Q., Klessig, D. F., He, S. Y. & Dong, X. Coronatine promotes *Pseudomonas syringae* virulence in plants by activating a signaling cascade that inhibits salicylic acid accumulation. *Cell Host Microbe* **11,** 587–596 (2012).

23. Du, L., Sanchez, C., Shen, B. Hybrid Peptide-Polyketide Natural Products: Biosynthesis and Prospects toward Engineering Novel Molecules. *Metab. Eng.* **3,** 78–95 (2000).

24. Bassett, R. A., Chain, E. B. & Corbett, K. Biosynthesis of ergotamine by *Claviceps purpurea* (Fr.) Tul. *Biochem. J.* **134,** 1–10 (1973).

25. Hedden, P. Gibberellin Biosynthesis. *eLS* (2001). doi:10.1002/9780470015902.a0023720

26. Brakhage, A. A., Spröte, M. T. P., Scharf, D. H., Al-Abdallah, Q., Wolke, S. M. & Hortschansky, P. Aspects on evolution of fungal beta-lactam biosynthesis gene clusters and recruitment of trans-acting factors. *Phytochemistry* **70,** 1801–11 (2009).

27. Brakhage, A. A & Schroeckh, V. Fungal secondary metabolites - strategies to activate silent gene clusters. *Fungal Genet. Biol.* **48,** 15–22 (2011).

28. Hertweck, C. The biosynthetic logic of polyketide diversity. *Angew. Chem. Int. Ed. Engl.* **48,** 4688–716 (2009).

29. Brakhage, A. A. Regulation of fungal secondary metabolism. *Nat. Rev. Microbiol.* **11,** 21–32 (2013).

30. Palmer, J. M. & Keller, N. P. Secondary metabolism in fungi: does chromosomal location matter? *Curr. Opin. Microbiol.* **13,** 431–6 (2010).

31. Shwab, E. K., Bok, J. W., Tribus, M., Galehr, J., Graessle, S. & Keller, N. P. Histone deacetylase activity regulates chemical diversity in *Aspergillus*. *Eukaryot. Cell* **6,** 1656–64 (2007).

32. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410,** 116–20 (2001).

33. Bayram, O. K. S. VelB / VeA / LaeA Complex Coordinates Light Signal with Fungal Development and Secondary Metabolism. *Science (80-.).* **320,** 1504–1506 (2008).

34. Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J. F. & Shilatifard, A. COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 12902–12907 (2001).

35. Marmorstein, R. Structure and function of histone acetyltransferases. *Cell. Mol. Life Sci.* **58,** 693–703 (2001).

36. Lee, Y., Kim, M., Han, J., Yeom, K-H., Lee, S., Baek, S. H. & Kim, V. N. MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23,** 4051–4060 (2004).

37. Palmer, J. M., Perrin, R. M., Dagenais, T. R. T. & Keller, N. P. H3K9 methylation regulates growth and development in *Aspergillus fumigatus*. *Eukaryot. Cell* **7,** 2052–60 (2008).

38. James, T. C. & Elgin, S. C. Identification of a Nonhistone Chromosomal Protein Associated with Heterochromatin in *Drosophila melanogaster* and Its Gene. *Mol. Cell. Biol.* **6,** 3862–3871 (1986).

39. Lomberk, G., Wallrath, L. & Urrutia, R. The Heterochromatin Protein 1 family. *Genome Biol.* **7,** 228 (2006).

40. Li, Y., Kirschmann, D. A. & Wallrath, L. L. Does heterochromatin protein 1 always follow code? *Proc. Natl. Acad. Sci. U. S. A.* **99,** 16462–9 (2002).

41. Platero, J. S., Hartnett, T. & Eissenberg, J. C. Functional analysis of the chromo domain of HP1. *EMBO J.* **14,** 3977–3986 (1995).

42. Jacobs, S. A., Taverna, S. D., Zhang, Y., Briggs, S. D., Li, J., Eissenburg, J. C., Allis, C. D. & Khorasanizadeh, S. Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3. *EMBO J.* **20,** 5232–5241 (2001).

43. Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C. & Kouzarides, T. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410,** 120–4 (2001).

44. Singh, P. B. & Georgatos, S. D. HP1: Facts, open questions, and speculation. in *J. Struct. Biol.* **140,** 10–16 (2002).

45. Blin, K., Medema, M. H., Kazempour, D., Fischbach, M. A., Breitling, R., Takano, E. & Weber, T. antiSMASH 2.0--a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* **41,** W204–12 (2013).

46. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14,** 178–92 (2013).

47. Palmer, J. M., Mallaredy, S., Perry, D. W., Sanchez, J. F., Theisen, J. M., Szewczyk, E., Oakley, B. R., Wang, C. C. C., Keller, N. P. & Mirabito, P. M. Telomere position effect is regulated by heterochromatin-associated proteins and NkuA in *Aspergillus nidulans*. *Microbiology* **156,** 3522–31 (2010).

48. Jacobs, S. A & Khorasanizadeh, S. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* **295,** 2080–3 (2002).

49. Ferri, N., Arnaboldi, L., Orlandi, A., Yokoyama, K., Gree, R., Granata, A., Hachem, A., Paoletti, R., Gelb, M. H. & Corsini, A. Effect of S(-) perillic acid on protein prenylation and arterial smooth muscle cell proliferation. *Biochem. Pharmacol.* **62,** 1637–1645 (2001).

50. Askham, L. Effective repellency concentration of bird shield repellent(TM)with methyl anthranilate to exclude ducks and geese from water

impoundments. in *Gt. Plains Wildl. Damage Control Work. Proceddings* 4 (1995). at <http://digitalcommons.unl.edu/gpwdcwp/421/>

51. Viñas, P., López Erroz, C. & Hernández Córdoba, M. Determination of methyl anthranilate and methyl N-methylanthranilate in beverages by liquid chromatography with fluorescence detection. *Chromatographia* **35,** 681–684 (1993).

52. Holtze, M., Mickiené, A., Atlas, A., Lindquist, L. & Schwieler, L. Elevated cerebrospinal fluid kynurenic acid levels in patients with tick-borne encephalitis. *J. Intern. Med.* **272,** 394–401 (2012).

53. Erhardt, S., Blennow, K., Nordin, C., Skogh, E., Lindström, L. H. & Engberg, G. Kynurenic acid levels are elevated in the cerebrospinal fluid of patients with schizophrenia. *Neurosci. Lett.* **313,** 96–98 (2001).

54. Levine, P. E. & N-j-, L. United States Patent [191. (1977).

55. State of California environmental protection agency office of environmental health hazard assessment safe drinking water and toxic enforcement act of 1986 chemicals known to the state to cause cancer or reproductive toxicity SEPTEMBER 11 , 2009 The Safe D. 1–19 (2009).

56. Hsieh, C.-H., Chao, K. S. C., Liao, H.-F. & Chen, Y.-J. Norcantharidin, Derivative of Cantharidin, for Cancer Stem Cells. *Evid. Based. Complement. Alternat. Med.* **2013,** 838651 (2013).

57. Wall, M. E., Wani, M. C., Cook, C. E., Palmer, K. H., McPhail, A. T. & Sim, G. A. Plant antitumor agents. I. Isolation and structure of camtothecin, a novel alkaloidal leukemia and tumor inhibitor from *Camptotheca acuminata*. *J. Am. Chem. Soc.* **88,** 3888–3890 (1966).

58. Thomas, C. J., Rahier, N. J. & Hecht, S. M. Camptothecin: Current perspectives. *Bioorganic Med. Chem.* **12,** 1585–1604 (2004).

59. Sun, Z. J., Chen, C., Sun, M. Z., Ai, C. H., Lu, X. L., Zheng, Y. F., Yang, B. F. & Dong, D. L. The application of poly (glycerol-sebacate) as biodegradable drug carrier. *Biomaterials* **30,** 5209–5214 (2009).

60. Pahl, H. L. & Baeuerle, P. A. A novel signal transduction pathway from the endoplasmic reticulum to the nucleus is mediated by transcription factor NF-kappa B. *EMBO J.* **14,** 2580–2588 (1995).

61. Kober, L., Zehe, C. & Bode, J. Development of a novel ER stress based selection system for the isolation of highly productive clones. *Biotechnol. Bioeng.* **109,** 2599–611 (2012).

62. Bush, K. *et al.* Tackling antibiotic resistance. *Nat. Rev. Microbiol.* **9,** 894–896 (2011).

63. Hancock, R. E. W. Collateral damage. *Nat. Biotechnol.* **32,** 66–68 (2014).

64. Morath, S. U., Hung, R. & Bennett, J. W. Fungal volatile organic compounds: A review with emphasis on their biotechnological potential. *Fungal Biol. Rev.* **26,** 73–83 (2012).

65. Helms, J. B. & Rothman, J. E. Inhibition by brefeldin A of a Golgi membrane enzyme that catalyses exchange of guanine nucleotide bound to ARF. *Nature* **360,** 352–354 (1992).

6. Appendices

6.1 List of primers used for cloning

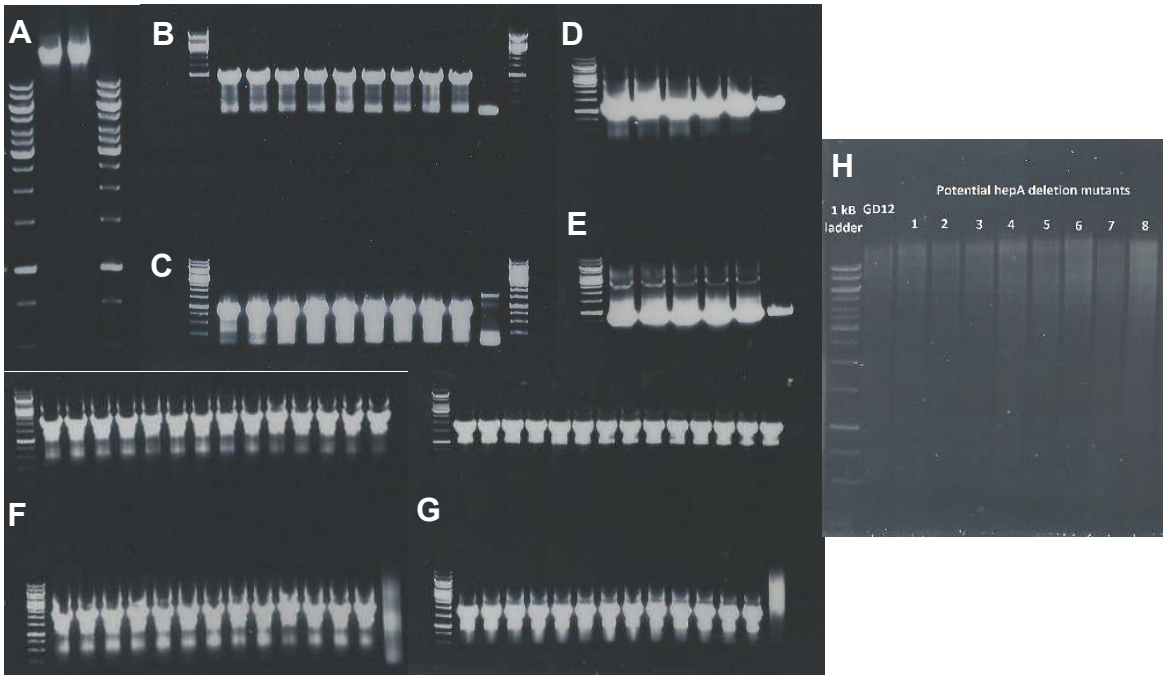
Primer	Sequence
hepA_LF-LP	5'-TGTAGACTGCTGCAGTGCACAA-3'
hepA_LF_LP2	5'-CACCTCGCACTGTATACTGGT-3'
hepA_LF-RP	5'-gtcgtgactgggaaaaccctggcgTGCAACGATGAGAAGCGATTGGT-3'
hepA_RF-LP	5'-tcctgtgtgaaattgttatccgctACTCAAGTGAGAAGACGTCGGA-3'
hepA_RF-RP	5'-ACACGATAAATGTGCCCGTCCT-3'
M13.LP	5'-TCCTGTGTGAAATTGTTATCCGCTGCCAGCATCCAA-3'
M13.RP	5'-GTCGTGACTGGGAAAACCCTGGCGGACGCCATACTC-3'
HY-split	5'-GGATGCCTCCGCTCGAAGTA-3'
YG-split	5'-CGTTGCAAGACCTGCCTGAA-3'

6.2 Conditions for polymerase chain reaction

	Product	Product size (bp)	Primer name	Template	Annealing temperature	Extension time
	hepA ORF	1118	hepA_ORF_LP hepA_ORF_RP	GD12 genomic DNA	52°C	1 min
1st round PCR	*hepA* LF	1016	hepA_LF_LP hepA_LF_RP	GD12 genomic DNA	52°C	1 min
	hepA RF	940	hepA_RF_LP hepA_RF_RP	GD12 genomic DNA	52°C	1 min
	HY	1200	M13_LP HY_split	*hph* cassette	58°C	1 min 10 sec
	YG	800	YG_split M13_RP	*hph cassette*	58°C	1 min 10 sec
2nd round PCR	*hepA* LF + HY	2216	hepA_LF_LP HY_split	*hepA* LF + HY	62°C	2 min 15 sec
	hepA RF + YG	1740	YG_split hepA_RF_RP	*hepA* RF + YG	62°C	1 min 50 sec

6.3 Gel images of PCR products and genomic digests.

All images are of 0.8 % TAE agarose gel, unless stated otherwise. All are against 10 µl Fermentas GeneRuler 1 Kb ladder, unless stated otherwise. (**A**) 5 µl GD12 genomic DNA, flanked by 5 µl of 1 kb ladder, run on a 1.2 % TAE agarose gel; (**B**) 9 x 25 µl LF PCR reactions, plus 1 x 25 µl H_2O control; (**C**) 9 x 25 µl RF PCR reactions, plus 1 x 25 µl H_2O control; (**D**) 5 x 25 µl HY PCR reactions, plus 1 x 25 µl H_2O control; (**E**) 5 x 25 µl YG PCR reactions, plus 1 x 25 µl H_2O control; (**F**) 27 x 25 µl LF+HY second round fusion PCR reactions, plus 1 x 25 µl H_2O control, (**G**) 27 x 25 µl RF+YG second round fusion PCR reactions, plus 1 x 25 µl H_2O control, (**H**) 20 µg of putative *hepA* deletion mutants, digested with restriction enzyme *Stu*I, run on a 1 % TBE agarose gel.



6.3 *T. hamatum* GD12 antiSMASH output

				TRICHODERMA ATROVIRIDE IMI206040		*ASPERGILLUS NIDULANS* FGSC_A4	
CLUSTER	**Type**	**Contig_Orf**	**Location**	**BLAST P (E-Value)**	**Annotation**	**BLAST P (E-Value)**	**Annotation**
1	**Nrps**	Ctg148_Orf000000	3048-46215	0	Non-Ribosomal Peptide Synthetase (B-Ketoacyl Synthase)	5E-169	Hypothetical Protein (Sidn3-Like) (Adenylation (A) Domain Of Siderophore-Synthesizing Nonribosomal Peptide Synthetases)
		Ctg148_Orf000001	46569-47957	0	Hypothetical Protein (Ppx/Gppa Phosphatase Family)	2E-118	Retrograde Regulation Protein 2 (AFU_Orthologue)(Ppx/Gppa Phosphatase Family)
		Ctg148_Orf000002	53737-55559	0	Hypothetical Protein (Lysp - Amino Acid Transporter And Metabolism)	0	Prnb (Proline-Specific Permease)
		Ctg148_Orf000003	56578-57148	5E-99	Hypothetical Protein (B-CA-Claded) Carbonic Anyhydrase	1.1	Hypothetical Protein (B-CA-Claded) Carbonic Anyhydrase
		Ctg148_Orf000004	61401-61949	1.8		3.6	
		Ctg148_Orf000005	64211-65303	0	Hypothetical Protein (Aldo-Keto Reductases (Akrs))	4e-139	Aflatoxin B1-Aldegyde Reducatse (Gli0-Like)
2	**Other**	Ctg207_Orf000000	3837-5107	4e-56	Non-Ribosomal Peptide Synthetase (Sidn3-Like) (Adenylation (A) Domain Of Siderophore-Synthesizing Nonribosomal Peptide Synthetases)	3e-50	Hypothetical Protein (Sidn3-Like) (Adenylation (A) Domain Of Siderophore-Synthesizing Nonribosomal Peptide Synthetases)
		Ctg207_Orf000001	12776-13278	6e-27	Non-Ribosomal Peptide Synthetase (Sidn3-Like) (Adenylation (A) Domain Of Siderophore-Synthesizing Nonribosomal Peptide Synthetases)	1e-23	Hypothetical Protein (Sidn3-Like) (Adenylation (A) Domain Of Siderophore-Synthesizing Nonribosomal Peptide Synthetases)

		Ctg207_Orf0002	13941-14134	7.1	Hypothetical Protein (SRR1)	2.7	Putative Zn(II)2Cys6 Transcription Factor (GAL4)
		Ctg207_Orf003	17639-17864	5.3	Hypothetical Protein (Mvim - Predicted Dehydrogenases And Related Proteins)	0.97	Hypothetical Protein (NAD Binding 8 - NAD(P)-Binding Rossmann-Like Domain)
		Ctg207_Orf04	18184-18402	2.7	Hypothetical Protein (Peptidase S24 S26 - Lexa/Signal Peptidase Superfamily)	2.3	Cytochrome P450, Putative
3		Ctg221_Orf000000	1386-2880	0.85	Hypothetical Protein (Transcription Factor Involved In Chromatin Remodeling- N-Acyltransferase Superfamily)	6.9	Hypothetical Protein (Polyletide Synthase Modules And Related Proteins)
		Ctg221_Orf00001	5833-6057	0.17	Hypothetical Protein	0.072	Hypothetical Protein (Predicted Acyl-Coa Transferases/Carnitine Dehydratase)
		Ctg221_Orf0002	6547-6805	7e-28	Hypothetical Protein (GAL4 - GAL4-Like Zncys6 Binuclear Cluster DNA-Binding Domain)	4.7	Hypothetical Protein (Anp1)
		Ctg221_Orf003	7369-8458	3e-67	Hypothetical Protein (Cytochrome P450)	6e-44	Hypothetical Protein (Cytochrome P450)
	Nrps	Ctg221_Orf04	11673-12667	0	Hypothetical Protein (Caic - Acyl-Coa Synthetases (AMP-Forming))	7e-12	Conserved Hypothetical Protein (AMP-Binding/Adenylate Forming Domain, Class I)
		Ctg221_Orf5	18912-19659	2e-41	Hypothetical Protein (Caic - Acyl-Coa Synthetases (AMP-Forming))	0.24	Hypothetical Protein (FHA - Forkhead Associated Domain)
		Ctg221_Orf6	19674-20349	7e-86	Hypothetical Protein (Condensation Domain)	6e-11	Hypothetical Protein (Condensation Domain)
		Ctg221_Orf7	18912-19659	7e-140	Hypothetical Protein (Prtases-Type I - Phosphoribosyl Transferase (PRT)-Type I Domain)	1e-97	Xanthine-Guanine Phosphoribosyl Transferase (Xpt1), Putative (PRT-Type I Domain)
		Ctg221_Orf8	19674-20349	9.9	Hypothetical Protein (Tht1-Like Nuclear Fusion Protein)	3.3	Hypothetical Protein (Kelch 5 Motif)

4	Other	Ctg235_Orf000000	2741-3428	5e-47	Non-Ribosomal Peptide Synthetase (Condensation Domain)	2e-48	Hypothetical Protein (Sidn3-Like) (Adenylation (A) Domain Of Siderophore-Synthesizing Nonribosomal Peptide Synthetases)
5	Other	Ctg322_Orf000000	2345-2733	8.3	Hypothetical Protein (IDO - Indoleamine 2,3-Dioxygenase)	2.3	Hypothetical Protein (Thif - Dinucleotide-Utilizing Enzymes Involved In Molybdopterin And Thiamine Biosynthesis Family 2)
		Ctg322_Orf00001	6729-7971	0	Polyketide Synthase (B-Ketoacyl Synthase)	6e-112	Hypothetical Protein (PKS - B-Ketoacyl Synthase)
		Ctg322_Orf0002	12589-15508	0	Hypothetical Protein (Caic - Acyl-Coa Synthetases (AMP-Forming))	0	Hypothetical Protein (Caic - Acyl-Coa Synthetases (AMP-Forming))
		Ctg322_Orf003	16812-18876	0	Hypothetical Protein (NAD Binding 8 - NAD(P)-Binding Rossmann-Like Domain)	2e-115	Conserved Hypothetical Protein (NAD Binding 8 - NAD(P)-Binding Rossmann-Like Domain)
		Ctg322_Orf04	19806-20249	9e-86	Hypothetical Protein (PT Ubia COQ2, 4-Hydroxybenzoate Polyprenyltransferase)	9e-18	Hypothetical Protein (PT Ubia COQ2, 4-Hydroxybenzoate Polyprenyltransferase)
6	Other	Ctg402_Orf000000	918-1451	1e-13	Putative Epoxide Hydrolase (Abhydrolase 6)	2.5	Hypothetical Protein (MFS - Major Facilitator Superfamily)
		Ctg402_Orf00001	1725-2867	0	Hypothetical Protein (ZIP Zinc Transporter)	7e-169	Hypothetical Protein (ZIP Zinc Transporter)
		Ctg402_Orf0002	5568-6952	1.4	Multidrug Resistance-Like Protein (MRP Assoc Pro)	0.62	Hypothetical Protein (Glycine Dehydrogenase; Provisional)
		Ctg402_Orf003	13102-14756	0	Hypothetical Protein (Lysp - Amino Acid Transport And Metabolism)	9e-180	Basic Amino Acid Transporter (Lysp - Amino Acid Transport And Metabolism)
		Ctg402_Orf04	16146-19994	0	Non-Ribosomal Peptide Synthetase (Adenylation Domain)	0	Hypothetical Protein (Adenylation Domain)
		Ctg402_Orf5	21469-23586	0	Hypothetical Protein (Arylsulfotransferase (ASST))	5e-54	Hypothetical Protein (Arylsulfotransferase (ASST))

		Ctg402_Orf6	24066-25549	1e-52	Hypothetical Protein (Rdrp - RNA Dependent RNA Polymerase)	0.094	Hypothetical Potein (Putative Methyltransferase)
		Ctg402_Orf7	28262-30976	0	Hypothetical Protein (GAL4 - GAL4-Like Zncys6 Binuclear Cluster DNA-Binding Domain)	5e-15	Putative Zn(II)2Cys6 Transcription Factor (GAL4)
		Ctg402_Orf8	31620-33356	0	Hypothetical Protein (RIO1 - Serine/Threonine Protein Kinase Involved In Cell Cycle Control)	0	Hypothetical Protein (RIO1 - Serine/Threonine Protein Kinase Involved In Cell Cycle Control)
		Ctg402_Orf9	33877-34175	0.34	Iron Sulfur Cluster Assembly Protein (Iscu-Like)	4.5	Conserved Hypothetical Protein (Protein Of Unknown Function)
		Ctg402_Orf10	34598-36026	0	Hypothetical Protein (Asp - Eukaryotic Aspartyl Protease)	1e-67	Hypothetical Protein (Asp - Eukaryotic Aspartyl Protease)
		Ctg402_Orf11	36550-37575	0	Hypothetical Protein (PCBER SDR A - Phenylcoumaran Benzylic Ether Reductase Like)	4e-63	Hypothetical Protein (NADB Rossmann)
		Ctg402_Orf12	37638-38198	4.0	Hypothetical Protein	0.66	Hypothetical Protein (NAD Binding 8 - NAD(P)-Binding Rossmann-Like Domain)
7	**T4pks-T1pks**	Ctg512_Orf0002	5984-7219	5e-71	Hypothetical Protein (Protein Of Unknown Function)	5e-17	Hypothetical Protein (Protein Of Unknown Function)
		Ctg512_Orf003	8817-10776	0	Hypothetical Protein (F-Box-Like)	0.082	Hypothetical Protein (F-Box-Like)
		Ctg512_Orf04	11191-13407	0	Glycoside Hydrolase Family 20 Protein (Glyco Hydro 20b)	6e-92	Hypothetical Protein (Glyco Hydro 20b)
		Ctg512_Orf5	17272-20424	0	Hypothetical Protein (Peptidases S8 5)	0.007	Hypothetical Protein (Peptidases S8 Protein Convertases Kexins Furin-Like)
		Ctg512_Orf6	23960-24555	1e-118	Hypothetical Protein (GFA - Glutathione-Dependent Formaldehyde-Activating Enzyme)	4e-70	Hypothetical Protein (GFA - Glutathione-Dependent Formaldehyde-Activating Enzyme)
		Ctg512_Orf7	24814-32496	0	Polyketide Synthase (Acyl Transferase Domain)	0	Hypothetical Protein (Acyltransferase Domain In PKS Enzymes)

		Ctg512_Orf8	34275-34924	1e-40	Hypothetical Protein (FSH1 - Serine Hydrolase)	2e-06	DUF341 Family Oxidoreductase, Putative (Serine Hydrolase (FSH1))
		Ctg512_Orf9	35114-35472	7.5	Hypothetical Protein (Aldo-Keto Reductases (Akrs))	1.9	Hypothetical Protein (Stkc Phototrophin-Like Protein)
		Ctg512_Orf10	36187-37811	0	Glycosyltransferase Family Protein 1	2e-76	UDP-Glucoronosyl And UDP-Glucosyl Transferase Family Protein
		Ctg512_Orf11	47529-48803	0	Hypothetical Protein (SWIRM Domain)	3e-46	SWIRM Domain Protein Fun19, Putative
		Ctg512_Orf12	51277-55040	0	Hypothetical Protein (Chromosomal Segregation Atpases)	1e-66	Hypothetical Protein (RIM-Binding Protein Of The Cytomatrix Active Zone)
8	**T1pks**	Ctg634_Orf000000	1330-1763	8e-68	Hypothetical Protein (Stress Responsive A/B Barrel Domain)	1e-27	Hypothetical Protein (Stress Responsive A/B Barrel Domain)
		Ctg634_Orf00001	2298-2252	1e-64	Hypothetical Protein (Uncharacterized Protein Containing Double-Stranded Beta Helix Domain)	1e-59	Hypothetical Protein (Uncharacterized Protein Containing Double-Stranded Beta Helix Domain)
		Ctg634_Orf0002	3327-4851	8e-81	Hypothetical Protein (MFS1 (Major Facilitator Superfamily)	4e-07	Hypothetical Protein (Ubih - 2-Polyprenyl-6-Methoxyphenol Hydoxylase And Related FAD-Dependent Oxidoreductases)
		Ctg634_Orf003	6543-7148	9.5	Hypothetical Protein (MOR2-PAG1 N)	3.7	Hypothetical Protein (CYCLIN)
		Ctg634_Orf04	9245-12169	0	Polyketide Synthase-Like Protein (B-Ketoacyl Synthase)	5e-172	Polyketide Synthase, Putative (Acyl Transferase Domain In Polyketide Synthase)
9	**Nrps**	Ctg665_Orf27	106307-108867	0	Hypothetical Protein (CNH Domain)	0	Hypothetical Protein (CNH Domain)
		Ctg665_Orf28	109276-117495	0	Hypothetical Protein (Multidrug Resistance Protein (Mdr1))	0	Hypothetical Protein (Multidrug Resistance Protein (Mdr1))
		Ctg665_Orf29	118280-120555	0	Hypothetical Protein (Trxb - Thioredoxin Reductase(Posttranslation Modification))	1e-119	Hypothetical Protein (Trxb - Thioredoxin Reductase(Posttranslation Modification))

		Ctg665_Orf30	121172-123777	2e-121	Hypothetical Protein (MFS1 (Major Facilitator Superfamily)	1e-58	Hypothetical Protein (MFS1 (Major Facilitator Superfamily)
		Ctg665_Orf31	125360-126796	0	Hypothetical Protein (Acetyltransferase Domain)	1e-114	Hypothetical Protein (Acetyltransferase Domain)
		Ctg665_Orf32	126998-129880	0	Long-Chain-Fatty-Acid-Coa Ligase	3e-125	Acyl Coa Synthetase
		Ctg665_Orf33	130614-136053	0	Hypothetical Protein (Sidn3-Like)(Adenylation (A) Domain Of Siderophore-Synthesizing Nonribosomal Peptide Synthetases)	6e-164	Hypothetical Protein (Sidn3-Like)(Adenylation (A) Domain Of Siderophore-Synthesizing Nonribosomal Peptide Synthetases)
		Ctg665_Orf34	136631-139527	5e-167	Hypothetical Protein (Protein Of Unknown Function)	3e-17	Hypothetical Protein (Protein Of Unknown Function)
		Ctg665_Orf35	144060-145115	0.39	Hypothetical Protein (Ribosomal P1 P2 L12p)	6.6	Hypothetical Protein (GET Complex Subunit GET2)
		Ctg665_Orf36	146362-147937	0	Hypothetical Protein (PAP-1-Like Conserved Region)	2e-13	Hypothetical Protein (PAP-1-Like Conserved Region)
		Ctg665_Orf37	148310-148453	0.18	Hypothetical Protein (Flavokinase)	0.20	Hypothetical Protein (Sulfur Sfnb)
		Ctg665_Orf38	149978-151129	0	Peroxisomal Membrane Anchor Domain-Containing Protein, Varient 1)	3e-46	Hypothetical Protein (Peroxisomal Membrane Anchor Protein (Pex14_N)
		Ctg665_Orf39	153788-153969				
		Ctg665_Orf40	154417-158953	0	Hypothetical Protein (BRX1)	4e-114	Ras Gtpase Similar To RAB11B
10		Ctg699_Orf000000	27-1806	0	Hypothetical Protein (Cytochrome P450)	4e-119	Hypothetical Protein (Cytochrome P450)
	T1pks	Ctg699_Orf00001	3302-4440	6e-140	Hypothetical Protein (Short Chain Dehydrogenase; Provisional)	5e-35	Hypothetical Protein (ARM)
		Ctg699_Orf0002	5084-5792	2e-143	Hypothetical Protein (Cupin 2)	2e-37	Hypothetical Protein (Predicted Acyl Esterases)
		Ctg699_Orf003	6296-7285	5e-164	Hypothetical Protein (Fabg - 3-Ketoacyl-(Acyl-Carrier-Protein) Reductase)	6e-27	Hypothetical Protein (OYE Like FMN)

	T1pks	Ctg699_Orf04	8929-17316	0	Polyketide Synthase-Like Protein (B-Ketoacyl Synthase)	2e-148	Hypothetical Protein (Acyl Transferase Domain In Polyketide Synthase)
		Ctg699_Orf5	30824-31471	2.4	Hypothetical Protein (RNA-Binding Protein Of The Puf Family, Translational Repressor)	0.042	Hypothetical Protein (Pumilio-Family RNA Binding Domain)
11		Ctg761_Orf000000	247-1023	3e-14	Hypothetical Protein (Mitochondrial Carrier Protein)	6e-08	Hypothetical Protein (Mitochondrial Carrier Protein)
		Ctg761_Orf00001	3383-5452	4e-13	Hypothetical Protein (Caic - Acyl-Coa Synthetases (AMP-Forming))	1e-15	Hypothetical Protein (Caic - Acyl-Coa Synthetases (AMP-Forming))
		Ctg761_Orf0002	6098-9032	9e-154	Hypothetical Protein (Amino Acid Kinase Family)	4e-180	Hypothetical Protein (Amino Acid Kinase Family)
		Ctg761_Orf003	9869-17713	6e-85	Hypothetical Protein (Acetyltransferase Domain)	9e-128	Hypothetical Protein (Acyl Transferase Domain In Polyketide Synthase Enzymes)
		Ctg761_Orf04	17874-18613	9e-44	Hypothetical Protein (B-Ketoacyl Synthase)	4e-51	Hypothetical Protein (Acyl Transferase Domain In Polyketide Synthase Enzymes)
		Ctg761_Orf5	20080-20796	4e-12	Hypothetical Protein (Mhpc - Predicted Hydrolases Or Acyltransferases)	3e-04	Hypothetical Protein (Polyketide Synthase Modules And Related Proteins)
		Ctg761_Orf6	21406-23296	4e-133	Hypothetical Protein (MFS1 (Major Facilitator Superfamily)	2e-13	Hypothetical Protein (MFS1 (Major Facilitator Superfamily)
		Ctg761_Orf7	24985-28770				
12	Bacteriocin	Ctg789_Orf13		0	Hypothetical Protein (PHD-Finger)	2e-21	Hypothetical Protein (PHD-Finger)
		Ctg789_Orf14	50669-52225	0	Hypothetical Protein (PA2G4-Like)	0	Hypothetical Protein (PA2G4-Like)
		Ctg789_Orf15	53823-56063	0	Hypothetical Protein (SNF5)	2e-172	Hypothetical Protein (SNF5)
		Ctg789_Orf16	57366-61955	3e-87	Hypothetical Protein	5e-32	Hypothetical Protein (TFIIF Alpha)
		Ctg789_Orf17	64196-65621	1e-156	Hypothetical Protein (Herpes BLLF1)	1.4	Hypothetical Protein (STAG Domain)
		Ctg789_Orf18	66237-67551				
13	Other	Ctg868_Orf000000	1174-4675	0	Hypothetical Protein (AFD Class	2e-44	Hypothetical Protein (AMP-

					I)		Binding Enzyme)
		Ctg868_Orf00001	8133-9623	4e-116	Hypothetical Protein (GAL4 - GAL4-Like Zncys6 Binuclear Cluster DNA-Binding Domain)	1.1	Hypothetical Protein (PX SNARE)
		Ctg868_Orf0002	11793-12482	3e-25	Hypothetical Protein (MDR7 - Medium Chain Dehydrogenase/Reductase)	2e-25	Zinc-Containing Alcohol Dehydrogenase, Putative (MDR7 - Medium Chain Dehydrogenase/Reductase)
		Ctg868_Orf003	17121-18004	1e-122	Hypothetical Protein (MFS1 (Major Facilitator Superfamily)	4e-38	Hypothetical Protein (MFS1 (Major Facilitator Superfamily)
		Ctg868_Orf04	18442-19760	0	Hypothetical Protein	1.9	Hypothetical Protein (FDH GDH Like)
		Ctg868_Orf5	20852-21557	5e-60	Putative Aspartate Aminotransferase (Aminotransferase Class I And II)	7e-67	Hypothetical Protein (Aminotransferase Class I And II)
		Ctg868_Orf6	22779-23226				
14		Ctg969_Orf0002	12853-14648	4e-74	Hypothetical Protein (Amidase)	5e-25	Hypothetical Protein (Amidase)
		Ctg969_Orf003	18781-19623	5e-169	Hypothetical Protein (ICL PEPM)	2e-52	Hypothetical Protein (ICL PEPM)
		Ctg969_Orf04	22003-23781	0	Hypothetical Protein (Peptidase S9)	2e-20	Hypothetical Protein (Peptidase S9)
	Terpene	Ctg969_Orf5	24394-27067	0	Oxidosqualene:Lanosterol Cyclase	0	Oxidosqualene:Lanosterol Cyclase
		Ctg969_Orf6	29430-30103	9e-126	Hypothetical Protein (Ank 2)	2e-21	Hypothetical Protein (Ank 2)
		Ctg969_Orf7	30149-31663	1.3	Hypothetical Protein (Ank 2)	0.19	Hypothetical Protein (Oxidase Reductase)
		Ctg969_Orf8	33173-34470	8e-146	Hypothetical Protein (FOG:L Zn-Finger)	3e-30	Hypothetical Protein (Fungal TF MHR)
		Ctg969_Orf9	38510-38518				
15		Ctg980_Orf16	46126-49113	0	Hypothetical Protein (MFS1 (Major Facilitator Superfamily)	1e-110	Hypothetical Protein (MFS1 (Major Facilitator Superfamily)
	Other	Ctg980_Orf17	49110-51175	2e-72	Hypothetical Protein (Verru Chthon Cassette Protein C)	9e-65	Putative Transcription Factor With C2H2 And Zn(2)-Cys(6) DNA Binding Domain
		Ctg980_Orf18	53354-55778	0	Hypothetical Protein (Trp-Synth-	2e-81	Metallopeptidase, Putative (Trp-

					Beta II)		Synth-Beta II)
		Ctg980_Orf19	55812-58393	0	Hypothetical Protein (MFS1 (Major Facilitator Superfamily)	0	Hypothetical Protein (MFS1 (Major Facilitator Superfamily)
		Ctg980_Orf20	58977-59437	5e-38	Hypothetical Protein (Dabb - Stress Responsive A/B Barrel Domain)	4e-13	Conserved Hypothetical Protein (Dabb - Stress Responsive A/B Barrel Domain)
		Ctg980_Orf21	60961-62168	0.069	Hypothetical Protein (Predicted Zn-Dependent Peptidases, Insulinase-Like)	0.092	Hypothetical Protein (AAA+)
		Ctg980_Orf22	66048-69958	0	Non-Ribosomal Peptide Synthetase (Adenylation Domain)	0	Hypothetical Protein (Condensation Domain)
		Ctg980_Orf23	70963-72763	0	Cytochrome P450	0	Cytochrome P450, Putative
		Ctg980_Orf24	73394-74831	0	Hypothetical Protein (Cytochrome B5-Like Heme/Steroid Binding Domain)	2e-166	Hypothetical Protein (Cytochrome B5-Like Heme/Steroid Binding Domain)
		Ctg980_Orf25	77197-79117	0	Hypothetical Protein (3-Hydroxybutyryl-Coa Dehydrogenase, Validated)	0	3-Hydroxybutyryl-Coa Dehydrogenase, Validated
		Ctg980_Orf26	79130-79297				
16		Ctg1001_Orf10	30521-30916	2e-67	Hypothetical Protein (Ribosomal P2)	2e-25	Hypothetical Protein (Ribosomal P2)
		Ctg1001_Orf11	42542-43485	0.001	Hypothetical Protein	7e-73	Hypothetical Protein
	T1pks	Ctg1001_Orf12	45194-45797	4e-92	Hypothetical Protein (Aldolase II)	1e-25	Hypothetical Protein (Aldolase II)
		Ctg1001_Orf13	46972-48409	6e-72	Acyl-Coa Synthetase (Caic)	2e-19	Hypothetical Protein (AFD Class I)
		Ctg1001_Orf14	50486-57860	0	Polyketide Synthase (Acyl Transferase Domain)	0	Hypothetical Protein (Acyl Transferase Domain)
		Ctg1001_Orf15	58336-62163				
17		Ctg1006_Orf000000	493-735	0.023	Hypothetical Protein (GAL4 - GAL4-Like Zncys6 Binuclear Cluster DNA-Binding Domain)	0.12	Hypothetical Protein (Glyco 32)
	Other	Ctg1006_Orf00001	3916-6711	0	Aminoadipate Reductase (Adenylation Domain Of NRPS)	1e-177	Hypothetical Protein (Condensation Domain)
		Ctg1006_Orf0002	7624-8809	2e-109	Hypothetical Protein (Cysteine	2e-37	Hypothetical Protein (Trp-Synth-

					Synthase)		Beta II)
		Ctg1006_Orf003	9601-9795				
		Ctg1006_Orf04	12907-13404	7e-12	Glutamate-Ammonia Ligase	2e-09	Hypothetical Protein (Glutamine Synthetase)
		Ctg1006_Orf5	14753-16283	0	Putative Acyl-Coa Dehydrogenase (Caia)	6e-97	Hypothetival Protein (Acyl-Coa Dehydrogenases)
		Ctg1006_Orf6	17532-20822	4e-116	Multidrug Resistance-Associated Protein	4e-12	Hypothetical Protein (CFTR Protein)
18	**Other**	Ctg1072_Orf000000	443-2798	2e-158	Non-Ribosomal Peptide Synthetase (Condensation Domain)	1e-52	Hypothetical Protein (Polyketide Synthase Modules And Related Proteins)
		Ctg1072_Orf00001	4154-5044				
19		Ctg1114_Orf000000	4480-8385	9e-173	Serine/Threonine Protein Kinase	9e-13	Hypothetical Protein (Protein Kinase, Catalytic Domain)
		Ctg1114_Orf00001	10224-11378	0	Hypothetical Protein (Pex24p)	2e-108	Hypothetical Protein (Pex24p)
	Terpene	Ctg1114_Orf0002	14765-16757	0	Geranylgeranyl Pyrophosphate Synthase (Trans IPPS HT)	7e-106	Hypothetical Protein (Geranylgeranyl Pyrophosphate Synthase (Trans IPPS HT))
		Ctg1114_Orf003	18039-19984	1e-164	Hypothetical Protein (CDC 14)	1e-70	Hypothetical Protein (CDC 14)
		Ctg1114_Orf04	22189-23724	0	Hypothetical Protein (IDO)	3e-151	Hypothetical Protein (IDO)
		Ctg1114_Orf5	25889-26646	0.86	Hypothetical Protein (Pleiotropic Drug Resistance)	5.8	Hypothetical Protein
20		Ctg1191_Orf000000	3129-3666				
		Ctg1191_Orf00001	6627-7780	1e-11	Hypothetical Protein (GAL4 - GAL4-Like Zncys6 Binuclear Cluster DNA-Binding Domain)	6e-05	Hypothetical Protein (GAL4 - GAL4-Like Zncys6 Binuclear Cluster DNA-Binding Domain)
		Ctg1191_Orf0002	9491-10045	1e-61	Hypothetical Protein (MFS1 (Major Facilitator Superfamily)	1e-42	Hypothetical Protein (MFS1 (Major Facilitator Superfamily)
	T1pks	Ctg1191_Orf003	12631-13057	2e-15	Hypothetical Protein (Short Chain Dehydrogenase, Validated)	2e-14	Putative Sterigmatocystin Biosynthesis Ketoreductase (Stce)
		Ctg1191_Orf04	18036-21122	9e-93	Hypothetical Protein (Adenylation Domain)	8e-42	Hypothetical Protein (Adenylation Domain)
		Ctg1191_Orf5	21778-23373	9e-68	Multidrug Resistance Protein (CFTR Protein)	8e-48	Hypothetical Protein (Multidrug Resistance Protein (Mdr1))

		Ctg1191_Orf6	28522-29759	0.97	Uncharacterised Protein (Mature Chain)	0.77	Hypotyhetical Protein (Urb2)
		Ctg1191_Orf7	30656-31123	7.2	Hypothetical Protein (Protein Kinase Domain)	4.5	Mitochondrial 3-Hydroxyisobutyryl-Coa Hydrolase, Putative
		Ctg1191_Orf8	33924-35571	9e-29	Hypothetical Protein (7-Keto-8-Aminopelargonate Synthetase And Related Enzymes	5e-46	Hypothetical Protein (Aspartate Aminotransferase (AAT)
		Ctg1191_Orf9	36880-37169	4.2	Cytochrome P450	4.6	Hypothetical Protein (MFS1 (Major Facilitator Superfamily)
		Ctg1191_Orf10	37399-44629	0	Polyketide Synthase (Acyl Transferase Domain)	0	Polyketide Synthase (Acyl Transferase Domain)
21	**Terpene**	Ctg1249_Orf000000	55-1427	0	Squalene Synthase (Trans-Isoprenyl Diphosphate Synthases, Head-To-Head)	3e-171	Farnesyl-Diphosphate Farnesyltransferase, Putative
		Ctg1249_Orf00001	3576-4397	4e-175	Hypothetical Protein (Nucleoside-Diphosphate-Sugar Epimerases)	2.4	CBF5 EMENI Centromere/Microtubule Binding Protein CBF5
		Ctg1249_Orf0002	4998-6675	0	Dnaj-Class Molecular Chaperone With C-Terminal Zn Finger Domain	8e-163	Dnaj-Class Molecular Chaperone With C-Terminal Zn Finger Domain
		Ctg1249_Orf003	8127-10565	6e-165	Hypothetical Protein (DNA Polymerase III Subunits Gamma And Tau; Provisional)	2e-36	Hypothetical Protein (Large Tegument Protein UL36, Provisional)
22	**Other**	Ctg1338_Orf000000	1391-2591	0	Hypothetical Protein (Cinnamyl-Alcohol Dehydrogenase Family Protein)	4e-24	Hypothetical Protein (Cinnamyl-Alcohol Dehydrogenase Family Protein)
		Ctg1338_Orf00001	11690-12264	1.5	Hypothetical Protein	2.8	DNA-Ligase
		Ctg1338_Orf0002	14830-17835	2e-98	Non-Ribosomal Peptide Synthetase (Phosphopantetheine Attachment Site)	2e-45	Hypothetical Protein (Adenylation Domain)
		Ctg1338_Orf003	19609-20789	7e-11	Non-Ribosomal Peptide Synthetase (Acyl Transferase Domain)	1e-10	Conserved Hypothetical Protein (Adenylation Domain)
		Ctg1338_Orf04	30925-31295	5.1	Hypothetical Protein (Protein	1.2	Hypothetical Protein (IKI3

					Kinases)		Family)
23	Terpene	Ctg1445_Orf7	23816-26716	0	Hypothetical Protein (Reca-Like Ntpases)	9e-129	Hypothetical Protein (DEAD-Like Helicases Superfamily)
		Ctg1445_Orf8	2949-28631	0	Hypothetical Protein (WD40 Domain)	0	Hypothetical Protein (WD40 Domain)
		Ctg1445_Orf9	34620-34998	4e-53	Hypothetical Protein (S-Adenosylmethionine-Dependent Methyltransferases, Class I)	9.9	Hypothetical Protein (Rpob - DNA-Directed RNA Polymerase)
		Ctg1445_Orf10	35213-35368	7e-10	Hypothetical Protein (A-Adenosylmethionine-Dependent Methyltransferases)	3.7	Hypothetical Protein (RNA-Binding Proteins)
		Ctg1445_Orf11	36130-37301	0	Terpene Synthase (Isoprenoid Biosynthesis Enzymes, Class I)	0.032	Hypothetical Protein (Isoprenoid Biosynthesis Enzymes, Class I)
		Ctg1445_Orf12	37588-38355	3e-158	Carbohydrate Esterase Family Protein 4	3e-72	Hypothetical Protein (Predicted Xylanase/Chitin Deacetylase)
		Ctg1445_Orf13	39112-40403	0	Hypothetical Protein (Cystathionine Beta-Lyases/Cystathionine Gamma-Synthases)	5e-151	Cystathionine Gamma-Synthase (Cystathionine Beta-Lyases/Cystathionine Gamma-Synthases)
24	Nrps	Ctg1455_Orf00001	8261-9292	0	Hypothetical Protein (NADB Rossmann)	4e-12	Isoflavone Reductase Family Protein
		Ctg1455_Orf0002	11028-12148	0	Hypothetical Protein (Cinnamyl-Alcohol Dehydrogenase Family Protein)	5e-97	Hypothetical Protein (Cinnamyl-Alcohol Dehydrogenase Family Protein)
		Ctg1455_Orf003	13064-14583	0	Hypothetical Protein (5beta-Reductase-Like Proteins)	1e-124	Hypothetical Protein (5beta-POR Like SDR A)
		Ctg1455_Orf04	19836-21352	0	Hypothetical Protein (NAD(P) +-Dependent Aldehyde Dehydrogenase Superfamily)	0	Hypothetical Protein (Aldehyde Dehydrogenase Family 2 Member)
		Ctg1455_Orf5	23583-25214	0	Hypothetical Protein (Choline Dehydrogenase And Related Flavoproteins)	0	Hypothetical Protein (Choline Dehydrogenase And Related Flavoproteins)
		Ctg1455_Orf6	25861-33183	0	Non-Ribosomal Peptide Synthetase (Adenylation Domain)	6e-57	Hypothetical Protein (Adenylation Forming Domain, Class I)

		Ctg1455_Orf7	42995-44778	0	Hypothetical Protein (L-Lysine 6-Monooxygenase)	9e-144	Hypothetical Protein (L-Lysine 6-Monooxygenase)
		Ctg1455_Orf8	47972-52330	0	Hypothetical Protein (GAL4 - GAL4-Like Zncys6 Binuclear Cluster DNA-Binding Domain)	6e-89	Hypothetical Protein (Nudix Hydrolase 7)
		Ctg1455_Orf9	52358-54104	0	Hypothetical Protein (Two Conserved Tryptophans Domain)	6e-129	Hypothetical Protein
25		Ctg1457_Orf000000	1287-2229	0	Hypothetical Protein (NADB Rossmann)	4e-60	Hypothetical Protein (Short Chain Dehydrogenase)
		Ctg1457_Orf00001	2935-4492	0	Hypothetical Protein (Pyoverdine/Dityrosine Biosynthesis Protein)	2e-61	Hypothetical Protein (Pyoverdine/Dityrosine Biosynthesis Protein)
		Ctg1457_Orf0002	6628-10438	0	Hypothetical Protein (Caic - Acyl-Coa Synthetases (AMP-Forming))	4e-82	Hypothetical Protein (AMP-Binding Enzyme)
		Ctg1457_Orf003	11009-11962	6e-158	Hypothetical Protein (NADB Rossmann)	3e-93	Hypothetical Protein (3-Ketoacyl-(Acyl-Carrier-Protein) Reductase)
	Other	Ctg1457_Orf04	13525-16203	2e-115	Hypothetical Protein (Glutathione S-Transferase [Posttranslational Modification, Protein Turnover, Chaperones]	2e-40	Glutathione Transferase, Putative
		Ctg1457_Orf5	17976-18353	2e-33	Hypothetical Protein (AAA-Like Domain)	9e-15	Conserved Hypothetical Protein
		Ctg1457_Orf6	22255-23689	0	Conserved Hypothetical Protein (2-Polyprenyl-6-Methoxyphenol Hydroxylase And Related FAD-Dependent Oxidoreductases)	9e-22	Hypothetical Protein (Salicylate Hydroxylase)
		Ctg1457_Orf7	24785-25728	0	Hypothetical Protein (Cysg - Siroheme Synthase)	2e-116	Siroheme Synthase Met8, Putative
		Ctg1457_Orf8	26323-27815	0	Hypothetical Protein (FYVE Domain; Zinc-Binding Domain)	2e-109	Hypothetical Protein (FYVE Domain; Zinc-Binding Domain)
		Ctg1457_Orf9	29341-31757	0	Hypothetical Protein (Kila-N Domain)	3e-157	Hypothetical Protein (Kila-N Domain)

26	Nrps	Ctg1464_Orf000000	537-2701	6e-50	Non-Ribosomal Peptide Synthease (NRPS Sidn3 Like (Adenylation Domain))	3e-48	Non-Ribosomal Peptide Synthease (NRPS Sidn3 Like (Adenylation Domain))
		Ctg1464_Orf00001	3338-3961	9e-32	Non-Ribosomal Peptide Synthetase (Phosphopantetheine Attachment Site)	4e-28	Hypothetical Protein (Adenylation Domain)
		Ctg1464_Orf0002	10332-11059	2e-28	Non-Ribosomal Peptide Synthetase (Adenylation Domain)	9e-29	Hypothetical Protein (Adenylation Domain)
		Ctg1464_Orf003	18662-19797	4e-96	Hypothetical Protein (GAL4 - GAL4-Like Zncys6 Binuclear Cluster DNA-Binding Domain)	0.43	Putative Zn(II)2Cys6 Transcription Factor (GAL4)
		Ctg1464_Orf04	21563-21633	0.026	Putative Stress Activated Mitogen Activated Protein Kinase Interacting Protein Sin1	0.006	Conserved Hypothetical Protein
27	T1pks	Ctg1540_Orf0002	10893-12821	8e-90	Hypothetical Protein (AAA+)	5e-37	Hypothetical Protein (AAA)
		Ctg1540_Orf003	13294-15194	3e-106	Non-Ribosomal Peptide Synthetase (Adenylation Domain)	3e-101	Hypothetical Protein (Adenylation Domain)
		Ctg1540_Orf04	20263-20882	2.1	Hypothetical Protein	5.1	Hypothetical Protein (Mitochondrial Carrier Protein)
		Ctg1540_Orf5	21085-22146	1.4	Hypothetical Protein (Type 1 Glutamine Amidotransferase (Gatase1)-Like Domain	0.085	Hypothetical Protein (GAL4 - GAL4-Like Zncys6 Binuclear Cluster DNA-Binding Domain)
		Ctg1540_Orf6	23323-24735	1e-172	Hypothetical Protein (Adenosylmethionine-8-Amino-7-Oxononanote Aminotransferase)	8e-122	Aminotransferase, Class III
		Ctg1540_Orf7	26046-32361	1e-53	Non-Ribosomal Peptide Synthetase (Adenylation Forming Domain)	4e-66	N-(5-Amino-5-Carboxypentanoyl)-L-Cysteinyl-D-Valine Synthase
		Ctg1540_Orf8	33167-34614	0	Hypothetical Protein (Nuf2)	1e-130	Nuf2
		Ctg1540_Orf9	38327-40524	0	Hypothetical Protein (Adaptin N)	0	Hypothetical Protein (Adaptin N)
		Ctg1540_Orf10	40905-41136				
28	T1 pk	Ctg1556_Orf0002	12866-13542	9e-75	Hypothetical Protein (GGCT-	1e-10	Hypothetical Protein (GGCT-Like

					Like Domain)		Domain)
		Ctg1556_Orf003	18280-19545	0	Hypothetical Protein (Glyco Hydro 61)	9e-69	Hypothetical Protein (Glyco Hydro 61)
		Ctg1556_Orf04	20412-21050	7.1	Hypothetical Protein (NADB Rossmann)	1.8	Hypothetical Protein (Ubih - 2-Polyprenyl-6-Methoxyphenol Hydroxylase And Related FAD-Dependent Oxidoreductases)
		Ctg1556_Orf5	21483-23846	3e-89	Hypothetical Protein (Protein Of Unknown Function)	6e-14	Hypothetical Protein (Protein Of Unknown Function)
		Ctg1556_Orf6	25752-27348	0	Hypothetical Protein (SMI1 KNR4)	3e-135	Hypothetical Protein (SMI1 KNR4)
		Ctg1556_Orf7	31459-31735	0.006	Hypothetical Protein (Metallo-Beta-Lactamase Superfamily)	7.0	N-(5-Amino-5-Carboxypentanoyl)-L-Cysteinyl-D-Valine Synthase
		Ctg1556_Orf8	33410-41280	0	Polyketide Synthase (Acyl Transferase Domain)	1e-149	Hypothetical Protein (Acyl Transferase Domain)
		Ctg1556_Orf9	42384-43810	8e-151	Hypothetical Protein (Cytochrome P450)	2e-52	Hypothetical Protein (Cytochrome P450)
29	**T1pks**	Ctg1612_Orf000000	400-7768	0	Polyketide Synthase (Acyl Transferase Domain)	3e-124	Hypothetical Protein (Acyl Transferase Domain)
		Ctg1612_Orf00001	8496-9407	4e-95	Hypothetical Protein (Znf C3H1)	1e-09	Hypothetical Protein (F0F1 ATP Synthase Subunit B)