Transcriptional reprogramming underpins enhanced plant growth promotion by the biocontrol fungus *Trichoderma hamatum* GD12 during antagonistic interactions with *Sclerotinia sclerotiorum* in soil

Sophie Shaw[1,2†], Kate Le Cocq[1,3†], Konrad Paszkiewicz[1], Karen Moore[1], Rebecca Winsbury[1,4], Marta de Torres Zabala[1], David Studholme[1], Deborah Salmon[1], Christopher R. Thornton[1], and Murray R. Grant[1]*‡

[1]Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter, UK

[2]Centre for Genome Enabled Biology and Medicine, University of Aberdeen, Aberdeen, UK

[3]Rothamsted Research, North Wyke, Okehampton, UK

[4]Department of Biological Chemistry, John Innes Centre, Norwich, UK

[†]The first two authors made an equal contribution to the work

*Correspondence: Murray Grant M.Grant@warwick.ac.uk

‡ Present address:

School of Life Sciences, Gibbet Hill Campus, University of Warwick,

Coventry, CV4 7AL, UK

Keywords: ***Trichoderma hamatum,*** **plant growth promotion, biocontrol,** ***Sclerotinia sclerotiorum,*** **RNA-seq**

Word count: 7133

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1111/mpp.12429

Abstract: 232
Introduction: 880
Results & Discussion: 3744
Concluding remarks: 299
Experimental Procedures: 938
Acknowledgements: 46
Figure, Table legends & Supplementary Table Figure Titles: 903

Summary

The free-living soil fungus *Trichoderma hamatum* strain GD12 is notable amongst *Trichoderma* strains in both controlling plant diseases and in stimulating plant growth, a property enhanced during its antagonistic interactions with pathogens in soil. These attributes, alongside its markedly expanded genome and proteome compared to other biocontrol and plant growth promoting *Trichoderma* strains, imply a rich potential for sustainable alternatives to synthetic pesticides and fertilisers for controlling plant disease and increasing yields. The purpose of this study was to investigate the transcriptional responses of GD12 underpinning its biocontrol and plant growth promotion capabilities during antagonistic interactions with the pathogen *Sclerotinia sclerotiorum* in soil. Using an extensive mRNA-seq study capturing different time points during the pathogen-antagonist interaction in soil, we show that dynamic and biphasic signatures in the GD12 transcriptome underpin its biocontrol and plant (lettuce) growth promotional activities. Functional predictions of differentially expressed genes demonstrate the enrichment of transcripts encoding proteins involved in transportation and oxidation-reduction reactions during both processes and an over-representation of siderophores. We identify a biphasic response during biocontrol characterised by a significant induction of transcripts encoding small-secreted cysteine rich proteins, secondary metabolite producing gene clusters and genes unique to GD12. These data support the hypothesis that *Sclerotinia* biocontrol

is mediated by the synthesis and secretion of antifungal compounds and that GD12's unique reservoir of uncharacterised genes is actively recruited during effective biological control of a plurivorous plant pathogen.

Introduction

There is ever increasing interest in novel solutions to enhancing crop yields. Agricultural intensification continues to be constrained by emerging, re-emerging and endemic pathogens that currently account for ~30% of losses in global crop production (Oerke and Dehne, 2004; Savary et al., 2012). Spiraling fertiliser costs and the associated financial carbon footprint of fossil fuel derived fertilisers means innovative solutions are needed to overcome sub-optimal soil fertility if marginal land is to be effectively farmed (Hillier et al., 2009). Similarly, growing public anxiety of the environmental and health impacts of synthetic chemical additives is restricting the numbers of effective pesticides (Carvalho, 2006). These factors are coupled to increasing resistance in fungicides and public distrust in genetic modification (Magnusson and Hursti, 2002; Frewer et al., 2004; Huesing et al., 2016). Therefore, it is essential to identify innovative approaches to enhance productivity. Harnessing the natural growth promotion and biocontrol properties of plant-beneficial rhizosphere microbes is an area with enormous potential to provide an increase in agricultural productivity while minimising inputs, waste and other negative impacts of agricultural intensification. To identify and exploit such biologics would require approaches that combine multi-scale studies of whole organism biology, genomics and chemistry.

The biological control and plant growth promotion (PGP) properties of *Trichoderma* species are well documented (Hermosa et al., 2012; Matarese et al., 2012; Omann et al., 2012; Ryder et al., 2012). Diseases are controlled through direct hyperparasitism of rhizosphere pathogens (Benitez et al., 2004), through competition (Ryder et al. 2012), through the production of antimicrobial compounds (Steindorff et al., 2014; Vos et al., 2014), through secretion of compounds that induce systemic resistance in plants (Lamdan et al., 2015) and in many cases most likely through a combination thereof (Howell, 2003). Additionally, certain *Trichoderma* strains are able to improve the productivity of monocot and dicot plants by accelerating seed germination and root and shoot development (Chang et al., 1986; Yedidia et al., 2001) and by increasing resilience to abiotic stresses such as drought (Donoso et al., 2008; Bae et al., 2009) and salinity (Mastouri et al., 2010; Brotman et al., 2013).

There is considerable agronomic interest in understanding the molecular mechanisms underpinning these beneficial processes, The transcriptional responses of three *Trichoderma* species with differing abilities to hyperparasitise *Rhizoctonia solani* are species-specific, with distinct gene families and discrete biochemical processes being deployed by each fungus, despite the presence of orthologs across all three biocontrol species (Atanasova et al., 2013). Thus biocontrol capabilities of different *Trichoderma* species appear heavily dependent on specific transcriptional reprogramming events during antagonistic interactions. For example, during PGP, *T. virens* cultured hydroponically with maize or tomato roots deployed both common and plant-specific transcriptional programmes, with important roles for glycosyl hydrolases and transporters implicated (Moran-Diez et al., 2015). Plant-pathogen-hyperparasite interactions induce specific gene expression (Steindorff et al., 2014), including over representation of transporters and carbohydrate active enzymes, similar to those

identified during *T. virens* PGP (Moran-Diez et al., 2015). Such robustness and flexibility in deploying different genome components to achieve biocontrol suggests different modes of pathogen perception feed into a convergent network. Collectively, these findings imply that both biocontrol and PGP can utilise common biological pathways to synthesise and deliver bioactive antimicrobial and PGP compounds. Indeed, initial transcriptional responses of *T. harzianum* grown in the presence of either tomato plants, chitin (an integral component of the fungal cell wall) or a simple glucose medium showed much greater similarity between plant and chitin treatments than glucose, inferring that during initial establishment common molecular mechanisms apply to the processes of PGP and biocontrol (Samolski et al., 2009).

The majority of *Trichoderma* species investigated to date usually exhibit either PGP or biocontrol properties. Notably, *T. hamatum* strain GD12 (herein after referred to as GD12) is a rare example that exhibits both properties (Ryder et al., 2012; Studholme et al., 2013). In lettuce microcosms containing low pH, nutrient-poor peat, GD12 elicits both biocontrol of root infecting pathogens and PGP (Ryder et al., 2012; Studholme et al., 2013). Genome sequencing of GD12 revealed a strikingly expanded genome with over 4500 predicted proteins without orthologs in four closely related *Trichoderma* species, including a large percentage of the predicted secretome - classically comprising small secreted cysteine rich proteins (SSCRPs). Some SSCRPs are likely to function as potential effector molecules that collaborate in biocontrol and PGP. Consistent with this, in other *Trichoderma* species SSCRPs have been shown to be up-regulated during PGP (Moran-Diez et al., 2015), during induction of systemic resistance (Lamdan et al., 2015) and during biocontrol (Atanasova et al., 2013). Thus, the expanded repertoire of SSCRPs unique to GD12 may contribute to its dual biocontrol and PGP properties.

In this study, we compared the transcriptional response of GD12 in peat during lettuce PGP to that induced during its biocontrol of the pathogen *Sclerotinia sclerotiorum* over 15 days, sampling at 6 time points. Using mRNA-seq, we captured the extent of transcriptional reprogramming in both GD12 and GD12-*S. sclerotiorum* mixed microcosms during the PGP and biocontrol processes and show that antagonistic interaction between *S. sclerotiorum* and GD12 activates unique and extensive transcriptional reprogramming early in the interaction, not seen under PGP conditions. Of particular note, underpinning these responses are specific subsets of both SSCRPs and predicted secondary metabolite associated gene clusters, indicative of important roles for both during PGP and disease amelioration.

Results & Discussion

Plant growth promotion of lettuce by *Trichoderma hamatum* strain GD12 is increased during antagonism of *Sclerotinia sclerotiorum*

In this study we investigated the transcriptional reprogramming in *Trichoderma hamatum* strain GD12 during its antagonism of the soil pathogen *Sclerotinia sclerotiorum.* GD12 biocontrol of *S. sclerotiorum* not only abrogates post-emergence pathogenesis of lettuce but, remarkably, further enhances lettuce growth promotion in comparison to GD12 alone (Studholme et al., 2013).

A peat microcosm time-course was set up in 6-well plates containing peat only (control), peat with GD12 only, peat with *S. sclerotiorum* only, or peat with GD12 and *S. sclerotiorum* – each well containing a single lettuce seedling. Plates were randomised daily. At 1 d, 2 d, 4 d, 7 d, 10 d and 15 d, random triplicate samples for each treatment - each sample comprising 2 pooled wells - were harvested and

immediately snap frozen in liquid nitrogen (see Fig. S1 for experimental design). Inoculation of peat with *Trichoderma hamatum* GD12 showed significant plant growth promotion (PGP) of lettuce across the 15-day time course compared to lettuce grown in peat alone (Fig. 1A). Additionally, antagonistic interactions following amendment with *Sclerotinia sclerotiorum* led to further enhancement of PGP (Fig. 1A) as reflected in the significant increases in dry weight in both roots and shoots during the antagonistic interaction compared to GD12 alone (Fig. 1B). As expected, no lettuce germinated in the peat microcosms containing *S. sclerotiorum* only.

Identification of differentially expressed genes differs significantly between DESeq2, CuffDiff2 and edgeR

To assess the transcriptional response of GD12 and *S. sclerotiorum* during plant-pathogen-hyperparasite interactions, RNA was extracted directly from the frozen microcosm samples. Bar-coded cDNA libraries were sequenced using Illumina 100 bp paired end reads. Resultant processed RNA reads were aligned to the GD12 reference genome (Studholme et al., 2013), and unmapped reads then aligned to the *S. sclerotiorum* reference isolate 1980 (Amselem et al., 2011). As expected, the majority of RNA reads aligned to the *T. hamatum* reference (Studholme et al., 2013) both in the GD12 only and mixed species microcosms (MSM) (Fig. 2). We hypothesise that the increased proportion of GD12 reads seen at 4-7 d in the MSMs is accounted for by stimulation of antimicrobial activity in the presence of *S. sclerotiorum*. Similarly, *S. sclerotiorum* formed the majority of the reads in the *S. sclerotiorum* only microcosms, and represented a declining proportion of the RNA population in the MSMs as the time course progressed, consistent with activation of GD12 biocontrol. At 1 d, *S. sclerotiorum* reads accounted for ~10-20% of total reads in the MSM, and this

had declined to less than 5% by 2 d. Notably, from 2 d, there was a much larger proportion of reads aligning to GD12 in the MSM than in the GD12 only microcosms, indicating enhanced transcriptional activity due to antagonism (Fig. 2). It should be noted that the increasing proportion of GD12 reads in the *S. sclerotiorum* only samples, notable at 15 d, is exacerbated by the alignment strategy which first aligned reads to the GD12 genome, then to the *S. sclerotiorum* genome, thus any *Sclerotinia* reads belonging to genes that are common with or highly similar to GD12 will first be identified as *Trichoderma* reads. Across all samples, on average 4% of reads aligned to both the GD12 and the *Sclerotinia* genome. Supplemental Table 1 contains full read statistics and read alignments to each specific genome.

All samples showed a proportion of RNA that did not align to either genome. This proportion increased over the duration of the experiment and most likely represents natural colonisation by airborne organisms in the controlled environment chamber over time. Notable was the increase in proportion of unknown reads over time, and in particular the larger proportion of these evident as early as 4 d in MSM microcosms, in which cryptic chemistries activated by GD12- *S. sclerotiorum* antagonistic interactions are present.

Of the 54 RNA extractions, the level of recovery from two replicates, 7 d GD12 replicate 3 and 10 d GD12 replicate 1, was insufficient for making high quality mRNA-seq libraries. On sequencing the 52 replicates, a single mixed species replicate at 10 d was excluded from further analysis due to an unusually high proportion of reads mapping to the *S. sclerotiorum* genome. This sample had been previously flagged, as it corresponded to a sample well that contained a single plant with atypically less growth (circled in Fig. S1), indicative of an isolated example of failed biocontrol. The

absence of these data points is accounted for by the analysis software and does not therefore have a negative effect on the dataset.

Our primary focus for this study was how *S. sclerotiorum* modified the transcriptional response of GD12 during antagonistic interactions in peat. The strategy was to compare expression profiles during this antagonistic response to those of GD12 alone which underpins PGP. Genes with significantly reduced expression or no expression in GD12 only microcosms would be assigned as up regulated in the mixed species microcosms and, *vice versa,* for those genes identified as up regulated in the GD12 only microcosms. Analysis of these differences is predicted to provide insight into how GD12 deploys its significantly expanded genome in biocontrol.

We first analysed the differences in the transcriptional response of GD12 and *S. sclerotiorum* using three separate differential expression analysis software packages. HTSeq (Anders et al., 2014) was used for quantification of reads at each predicted gene location, which were then used for DESeq2 (Love et al., 2014) and edgeR (Robinson et al., 2010) analyses. In addition, Cufflinks (Trapnell et al., 2010) was used to quantify transcripts, which were then analysed by CuffDiff2 (Trapnell et al., 2013). All three methods used a significance threshold of <0.05 for either the adjusted p value or q value, representing a 5% false discovery rate, corrected for multiple testing using the Benjamini Hochberg method (Benjamini and Hochberg, 1995).

The differentially expressed genes (DEGs) relative to each condition differed strikingly depending on the software chosen (Fig. 3). This appears to be a consequence of the different statistical models used between software packages. CuffDiff2 applies a negative binomial model followed by computation of a t-like statistic based upon fragments per kilobase per million mapped read (FPKM) values (Trapnell et al., 2010).

DESeq2 and edgeR both use raw count data. However, DESeq2 uses a negative binomial model with a general linear model for comparison (Love et al., 2014) whereas edgeR uses a similar over-dispersed Poisson model to explain variation followed by an Empirical Bayes model (Robinson et al., 2010). It is surprising that such remarkable differences are generally not considered when undertaking mRNA-seq studies, as these statistical software have been compared previously using data from the well annotated *Saccharomyces cerevisiae* genome (Nookaew et al., 2012). Like the results seen in Fig. 3, the Nookaew study showed large variations in the numbers of differentially expressed genes; edgeR identified a much higher number of genes, whereas CuffDiff2 identified the largest number of DEGs. Despite these differences, the overall biological interpretations based upon gene ontology (GO) did not differ between software programmes (Nookaew et al., 2012).

Log-fold changes as calculated by DESeq2 and edgeR can easily be compared because the gene model used by each programme is identical. In our data, these log-fold changes showed highly significant correlations (Pearson correlation >0.97) despite differences in the number of DEGs identified (Fig. S2). Comparisons with CuffDiff2 are more complex as it only uses gene models as a guide and exact locations of transcripts are predicted *de novo*. In addition to our 12591 gene predictions, Cufflinks predicted a further 8076 transcripts and frequently altered gene coordinates thus contributing to the differences observed in Fig. 3.

For all subsequent analyses we chose to use DESeq2 for the following reasons. First, CuffDiff2 predicts *de novo* transcript coordinates that do not match the FGenesH predictions in the GD12 reference (Studholme et al., 2013), notably the bespoke predictions of small secreted cysteine rich protein (SSCRP) encoding genes associated with the GD12 secretome. Secondly, despite finding similar biological

ontologies between DEGs (Nookaew et al., 2012) it has been reported that edgeR produces a higher number of false positives (Zhang et al., 2014). To use the intersection of all three programmes would multiply all biases, thus based upon the arguments above the more conservative DESeq2 was used to enable a focus on the most robustly identified differential signals.

T. hamatum GD12 has a distinct transcriptional profile when confronted with *S. sclerotiorum*

We defined "expressed" as having a DESeq2 normalised count of ≥ 25. Using this criteria we identified ~50% of the 12591 predicted genes in GD12 and MSMs treatments as being expressed during the time course. Figure 4 summarises the differences in expression between GD12 and GD12/*S. sclerotiorum* peat amendments over the 15-day time-course. The core shared transcriptome between treatments comprised between ~4,300 to ~5,800 genes. A greater percentage of DEGs were present in GD12 treatments between 1-4 d (Fig. 4). By contrast, between 7-15 d the MSM treatment had a larger proportion of DEGs, numerically representing nearly 25% of the core transcriptome at 10 d (Fig. 4). The *S. sclerotiorum* transcriptional profile was not the primary objective of this study. We initially undertook analysis of differential genes expression in *S. sclerotiorum,* however the dramatically reduced number of reads in the MSMs compared to the monoculture - due to aggressive biocontrol by GD12 - meant that any differential expression analysis of the *S. sclerotiorum* transcriptome would not be statistically valid.. The summary of read numbers and raw data are available (Table S1; GEO accession number GSE67909).

DEGs identified between GD12 only and MSMs and their annotations (Blast2GO predictions (Conesa et al., 2005)) are summarised in Table 1 and catalogued in Supplementary Materials 1-4. The transcriptional reprogramming is primarily the result of differential expression rather than induction of novel transcripts, as illustrated in Fig. 5A, which captures the strikingly sharp increase in expression between 1 d and 2 d following GD12 amendment. Previous studies conducted in soil-less systems over short time periods have shown early transcriptional reprogramming of *T. harzianum* during the initial stages of biocontrol (Samolski et al., 2009; Steindorff et al., 2014). By contrast, our experiments conducted in more complex peat microcosms and over a more sustained experimental period reveal a biphasic dynamic with a further significant number of DEGs evident in MSMs at 10 d and 15 d (Fig. 5A) when differential PGP between treatments was evident (Fig. 1). This later transcriptional response may be induction of components associated with PGP or possibly a biocontrol response to the establishment of natural communities, as inferred from unmapped reads in Fig 2. As illustrated above, the majority of DEGs were present under both conditions. Only 7 genes were uniquely expressed in GD12 only, whereas more than 15 times as many unique genes (114) were induced in MSMs over the 15-day time course (Table 1 and Fig. 5A). Predicted gene annotations are provided in Supplementary Materials 3 & 4; determining the function of these genes may facilitate better understanding of biocontrol mechanisms.

Hierarchical clustering showed that the transcriptional response within and across treatments over the 15 days was remarkably dynamic (Fig. 5B), with only nine genes in GD12 only microcosms and 22 genes during antagonism challenge – of ~6000 expressed genes - significantly up regulated at all time points (Table 2). These nine GD12-specific genes encode, amongst others, methyltransferases, a malic acid

transporter and a potassium uptake protein (Table 3). The predicted function of the 22 genes which were persistently up-regulated in the MSM encompasses both components of antagonism and those required for a protective strategy to restrict self-damage to GD12 occurring from antimicrobial compounds released during antagonism. Four are predicted to encode transporters, including one associated with resistance to the antifungal triazole, fluconazole. Two genes are predicted to encode dienelactone hydrolases, involved in degradation of catechol by bacteria (Schmidt and Knackmuss, 1980; Schlömann et al., 1993), and a pyoverdine dityrosine biosynthesis gene. In bacteria, pyoverdine acts as iron siderophore, enabling iron sequestration from the environment (Meyer, 2000). Notably, iron binding siderophores have been shown to be important for the saprotrophic lifestyle of *Trichoderma* species (Anke et al., 1991; Renshaw et al., 2002; Hoyos-Carvajal et al., 2009). Additionally, of three *Streptomyces* isolates producing siderophores tested for their PGP capabilities, isolates producing siderophores were found to have the greatest increase in plant biomass (Verma et al., 2011).

The potential functional roles of DEGs were investigated using gene ontology (GO) terms. While both transport and oxidation-reduction (redox) processes terms were strongly over-represented across all time points under both conditions (Fig. 5C), redox processes were significantly more abundant in the MSMs at 1 d, 7 d and 10 d. Strikingly, at 1 d there were twice as many redox associated DEGs in the MSM compared to GD12 only. This suggests that redox components are actively recruited during early antagonistic events. There were also contrasting patterns in “transport” process DEGs, with a marked over representation in the GD12 only microcosms at 2 d, whereas this signature was not observed until 10 d in MSMs (Fig. 5C), suggestive

of initial active prioritisation of *S. sclerotiorum* biocontrol over PGP observed at 10 d (Fig. S1).

qPCR validation of early differentially expressed genes

Given the gross differences seen in the different algorithms for predicting DEGs, we validated the DESeq2 predicted expression profiles of selected genes at 2 d, or 2 d and 15 d by qRT-PCR (see Table S2 for primers) based upon their predicted profiles in the two treatments. Given our interest in early events we assayed the following patterns; early induction in mixed species microcosms (ANCB01014346.1:6322-7713; Fig. S3A), a *SSCRP* with sustained induction in mixed species microcosms (ANCB01009946.1:717-1539; Fig. S3B), an unknown transcript showing early induction in GD12 only microcosms (ANCB01002344.1:7484-8547; Fig. S3C) and a *SSCRP* with early induction in mixed species microcosms (ANCB01002310.1:194-864; Fig. S3D). In summary, these assays confirmed the early response expression profiles observed in the RNA sequencing data.

***T. hamatum* GD12 – *S. sclerotiorum* interactions elicit a biphasic transcriptional response**

We next focused on features that distinguished the antagonistic interaction using hierarchical clustering and identified two distinct sets of genes up-regulated by GD12 in response to *S. sclerotiorum* at early (1 d to 4 d) and late (4 d to 15 d) time points (Fig. 6). Eighteen genes showed higher expression during antagonism at early time points, two of which are hypothesised to have roles in the breakdown of the fungal cell wall and membrane. These two genes are orthologs of *T. hamatum* strain LU593 which were previously shown to be up-regulated during confrontation with *S.*

sclerotiorum; the alkaline proteinase *prb1* (Steyaert et al., 2004) and a metalloendopeptidase (Carpenter et al., 2005). These studies thus further validate our data, demonstrating commonality in *Trichoderma* transcriptional responses to biocontrol function, and supports our prediction that this cluster defines a new set of antagonism specific genes. Notably, within the cluster, seven of the 18 genes encode strong homology to major facilitator superfamily (MFS) transporters and oligopeptide transporters. Given transporters prominent role in active mobilisation of secondary metabolites involved in antagonism (and/or roles in protection of the antagonist), the unexpectedly small number (Fig. 5C) but pronounced expression levels, of transporter genes associated with biocontrol at early time points is indicative of important functional roles early in antagonistic interactions with *Sclerotinia.* Other genes in this co-expressed cluster included an oxidoreductase, a methyltransferase, two short chain dehydrogenase genes, a glycoside hydrolase, a HSCARG dehydrogenase (involved in NADPH sensing (Zhao et al., 2008) and regulation of H2A ubiquitination (Hu et al., 2014)) or chaperone proteins, all implicated in modification of metabolites or gene regulation. The fact that 16 of 18 genes in the MSM early response cluster have predicted functions (enzymatic and transport) implies modification of small molecules/biosynthetic enzymes is a key early event in establishment of GD12 biocontrol capability.

In addition to this early transcriptional response associated with active antagonism of *S. sclerotiorum* (Fig. 1), a second phase of transcriptional reprogramming in MSMs relative to GD12 only microcosms was evident from 4 d to 15 d that may contribute to the enhanced PGP observed experimentally (Fig. 1). Six of the 103 genes up-regulated during the later time points in MSMs were unique to GD12 and a further 23 encoded hypothetical *Trichoderma* specific proteins, suggesting that antagonism

recruits genes capable of synthesis of novel bioactive compounds. For 74 of the 103 genes up-regulated late in MSMs a functional inference was possible, revealing recruitment of quite different complements of genes between the early and late phases of GD12 antagonistic interactions. Notable were genes with prominent roles in metal acquisition and carbon utilisation, indicative of saprotrophic colonisation or rhizosphere interactions. These genes include a PiT family inorganic phosphate transporter, four copper transporters, five hydrolases, a gene predicted to enable unusual carbon/nitrogen source utilisation and three iron siderophore transporters. The latter reinforces the importance of siderophores not only to the saprotrophic lifestyle but also to biocontrol and rhizosphere colonisation. Siderophore production has previously been linked to the biocontrol and PGP capabilities of the bacterium *Bacillus subtilis* CAS15 (Yu et al., 2011), however this study is the first to implicate the importance of siderophores in fungal mediated biocontrol.

At the later time points in the MSMs, a number of genes whose products are associated with protective functions were induced. These genes encoded five MFS transporters, one fluconazole resistance protein, two trichothecene 3-O-acetyltransferases which have been shown to protect against mycotoxins (Kimura et al., 1998) and two nitric oxide and mycotoxin protective dioxygenases. Additionally, genes with potential functional roles in the production of lipids (essential precursors in secondary metabolite biosynthesis) and production of antifungals were strongly over-represented including a squalene epoxidase involved in sterol biosynthesis, whose expression in *T. harzianum* is triggered by exposure to *Penicillium digitatum* cell wall extracts (Cardoza et al., 2006). Furthermore, genes encoding a perilipin-like protein involved in lipid accumulation, a polyprenol reductase, a phenazine biosynthesis protein, a SUR2-type hydroxylase/desaturase with a role in fatty acid biosynthesis, and

a gene with 53% identity to a gliotoxin biosynthetic gene, were also over-represented. So collectively, late transcriptional reprogramming in MSMs appears strongly geared towards production of secondary metabolites with protective or antifungal potential.

Small secreted cysteine rich proteins expressed during PGP and biocontrol

Small secreted cysteine rich proteins (SSCRPs) are predicted effectors hypothesised to have important roles in *Trichoderma* mediated rhizosphere interactions (Moran-Diez et al., 2015), induction of systemic resistance (Lamdan et al., 2015) and antagonism (Atanasova et al., 2013). We identified 153 putative SSCRPs in the GD12 genome (Studholme et al., 2013) of which more than half (79) were expressed in GD12 only and MSMs. Notably, of the 55 SSCRPs significantly up-regulated in the GD12 only microcosms (Table S3), the majority (~60%) were expressed at 4 d or earlier. Sixteen were GD12 “unique” (no annotation) and 14 represented “hypothetical” proteins. Using WoLF PSORT to predict cellular location, the majority of early GD12 only induced SSCRPs were predicted as mitochondrial localised (~30%; Fig. 7A), with 13% extracellular and 15% plasma membrane localised. The role of these genes as effectors is currently under investigation.

In MSMs 43 putative SSCRP encoding transcripts were significantly up regulated, of which, 50% had no predicted annotation (Table S4). In contrast to the GD12-only SSCRPs which were primarily induced 4 d or earlier, MSM SSCRPs were induced at later stages (4 d onwards, Fig. 7B) underlining the selective transcriptional dynamics under the two different conditions. Furthermore, WoLF PSORT predicted that ~50% of MSM induced SSCRP genes encoded components likely to be extracellular or plasma membrane localised (Fig. 7A). Collectively, these data demonstrate that a number of

GD12 gene encoding unique SSCRPs are differentially deployed under antagonistic interactions and are predicted to function in the immediate environment.

antiSMASH analysis reveals potential for secondary metabolite production in *T. hamatum* GD12

Novel secondary metabolite production, most notably from non-ribosomal peptide synthetases (NRPS) and polyketide synthases (PKS), are predicted to be important for both *Trichoderma* biocontrol and PGP (Baker et al., 2012; Mukherjee et al., 2012). Using antiSMASH (Blin et al., 2013) we identified 32 gene clusters in GD12 including NRPSs, type 1 PKSs and terpene synthases (Table S5). Strikingly, only five antiSMASH clusters were homologous to other species further highlighting the capacity for novel metabolite production by GD12. Most surprising was the prediction that cluster 31 homologues were bacterial specific.

We validated that antiSMASH predictions correlated to FGenesH predicted GD12 gene locations and mRNA-seq reads, thus highlighting secondary metabolite biosynthetic gene clusters potentially linked to PGP and biocontrol. The cluster at ANCB01004730.1 contains a single gene, with BlastP hits to a polyketide synthase (β-ketoacyl synthase), which shows increased expression at 4 d in GD12 only (Fig. 8A). The contig ANCB01008794.1 contains a two-gene cluster with homology to a NRPS from *T. atroviride* (the closest relative of *T. hamatum* for which complete genome sequence data are available) one of which showed significantly increased expression in the MSMs on 2 d (Fig. 8B). Transcripts from the predicted terpene synthase cluster at ANCB01010770.1 were significantly more abundant in GD12 only microcosms on 2 d (Fig. 8C). AntiSMASH predicted a two gene NRPS cluster in contig ANCB011924.1, whereas FGenesH predicted a single gene. In GD12 only, this gene

showed significantly higher expression on 2 d whereas in the mixed species its expression is significant on 15 d (Fig. 8D), in line with our biphasic prediction. The long chain fatty acid ligase encoded in the NRPS two gene cluster predicted in contig ANCB01005027.1 showed significant increase in transcription in MSMs on 4 d, 7 d and 10 d (Fig. 8E). ANCB01007958.1, predicted to contain a type 1 PKS, showed increased expression in GD12 only on 2 d and in the mixed species on 7 d and 10 d (Fig. 8F). Most notable, the NRPS containing contig ANCB01007961.1, predicted to encode aerothricin synthetase, showed significantly higher expression in MSMs on 1 d and 10 d (Fig. 8G). Aerothricin synthetase is involved in the synthesis of the novel cyclic antifungal aerothricins, which inhibit the β-1,3-D-glucan component of fungal cell walls (Pharmaceutica, 2002). AntiSMASH predicted a three-gene PKS cluster on ANCB01010115.1 (FGenesH predicted a single gene) with significantly higher reads in the mixed species microcosms on 15 d (Fig. 8H). This antiSMASH analysis reinforces specific transcriptional differences between treatments in potential antimicrobial encoding secondary metabolism clusters and in three instances demonstrates late induction of both polyketide synthase (ANCB01007958.1, ANCB01010115.1) and NPRS clusters (ANCB011924.1) indicating novel secondary metabolism is part of the second biphasic response seen in MSMs during elaboration of biocontrol and PGP (Fig. 6).

Concluding remarks

This study was designed to capture the transcriptional reprogramming associated with the dual PGP and biocontrol uniquely exhibited by *Trichoderma hamatum* GD12 (Ryder et al., 2012; Studholme et al., 2013) in peat microcosms. GD12 is predicted to encode more than 40% unique proteins compared to its most closely related

sequenced relative, *T. hamatum*. With ~50% of the transcriptome being expressed in the 6 samples that covered a 15 day period, we were able to capture the underlying transcriptional reconfiguration that is associated with functionally known and novel transcripts that are recruited to deploy the diverse armory of cryptic chemistries that are potentially encoded in GD12 genome. We identified biphasic response signatures and were able to assign specific transcriptional dynamics, and often transcripts of unknown function, to GD12 biocontrol and PGP strategies. Comparing the GD12 "PGP transcriptome" with that of its earlier inferred "biocontrol transcriptome" has allowed us to identify new sets of antagonism specific genes and their temporal dynamics. Functional analysis demonstrates that key classes of genes are up regulated during lettuce growth promotion and during antagonistic interactions with the root-infecting pathogen *S. sclerotiorum*, including transporter and oxidation-reduction genes, with a role for siderophores being prominent. We further identified the suite of SSCRP transcripts expressed during antagonism and the transcripts associated with secondary metabolism clusters are predicted to be involved in the synthesis of a collection of antimicrobials. Collectively, our data indicate that the unique genomic potential of GD12 is deployed as an integral part of the establishment and maintenance of biocontrol of a plant pathogen with wide host range. This study provides the foundation to explore the functional contribution of this transcriptional reprogramming, particularly the novel GD12 components predicted to be involved in secondary metabolite formation, in inter-species communication during the processes of PGP and plant disease control.

Experimental procedures

Fungal strains *Trichoderma hamatum* strain GD12 (GenBank accession no. AY247559 (Ryder et al., 2012)) and *S. sclerotiorum* strain M448 (courtesy of Jon West, Rothamsted Research) were both cultured and maintained on Potato Dextrose agar at 26 ˚C under a 16 h light regime.

Plant growth assays Lettuce biocontrol assays were carried out essentially as described (Ryder et al., 2012; Studholme et al., 2013) with the following amendments. Sphagnum moss peat (Shamrock) was supplemented with $CaCO_3$ (Sigma) at 0.5 g/L to adjust the pH to ~5.8, and assays were carried out in 6-well plates (Nunc) containing 50 g peat per plate. Four peat microcosm treatments (illustrated in Fig. S1) were established as follows; (i) peat only (control), (ii) peat amended with *T. hamatum* GD12 inoculum alone (GD12 only), (iii) peat amended with *S. sclerotiorum* inoculum alone (*S. sclerotiorum* only), or (iv) peat amended with a mixture of both species (GD12 + *S. sclerotiorum*). Each treatment consisted of six 6 well plates, totaling 36 wells per treatment, of which two were pooled into one sample at each time point. Two lettuce (*Lactuca sativa* cultivar Webb's Wonderful) seeds were sown in each well and the microcosms placed in a fully randomized design at 24 °C with a relative humidity of 90% and a 16 h fluorescent light regime. The first emerging lettuce plant in each well was retained while the other was removed. At six time points (1 d, 2 d, 4 d, 7 d, 10 d and 15 d), six wells were selected at random from each treatment, above ground lettuce foliage was removed and the peat containing lettuce root and fungal biomass from two plugs were pooled. Each sample was immediately snap frozen in liquid nitrogen and stored at -80 °C prior to RNA extraction.

For dry weights a replicate experiment was carried out as above. After 14 days, plants were harvested, washed and oven dried (75 ^{o}C). Shoot and root weights of dried material were determined and the data analysed by using ANOVA and Tukey Post hoc tests.

Fungal RNA extraction from peat samples Peat was homogenised to a granular powder using a Blender (Waring) in the presence of liquid nitrogen. From this material, a 2 g aliquot of peat was used to extract total RNA using the PowerSoil® Total RNA Isolation kit (Mo Bio, Carlsbad, CA) following the manufacturer's instructions. All 72 soil microcosm samples were quantified by both RNA gels and BioAnalyser traces (Agilent) to assess quality. For samples falling below the threshold criteria, new RNA extractions were performed on the homogenised material. RNA from Peat only microcosms were excluded from further analysis as the quantity of RNA extracted was negligible. Analysis proceeded on the remaining 54 samples.

Library preparation and sequencing Two-µl of diluted ERCC RNA control spike-in mix 1 or control spike-in mix 2 (Life Technologies) was added to 1 µg of soil RNA and the volume adjusted to 50 µl with nuclease free water. Libraries were prepared using the TruSeq RNA application from Perkin Elmer for Illumina TruSeq RNA library preparation on a Sciclone NGS platform. Library quality and quantity was determined by LabChipGX high sensitivity assay (Perkin Elmer). 100-bp paired end sequencing was then carried out using an Illumina HiSeq2500 platform with v3 reagents.

qPCR validation of differentially expressed genes Total RNA (1 µg) was DNAse treated with RNAse-free DNAse (Ambion, Life Technologies) and subsequently first

strand cDNA synthesis was carried out with Superscript III First Strand SuperMix (Invitrogen, Life Technologies) according to the manufacturers instructions. Real-time PCR was performed with the SensiFAST™ SYBR kit (Bioline) in a Corbett Rotor-Gene 6000 (Corbett). In all the cases, three technical replicates were run for each cDNA sample. Supplementary Table 2 reports primer combinations and Figure S3 validation results.

Data analysis The following Illumina pipeline processing was performed. Raw reads of low quality (phred score < 30) and adapter sequences were removed using Fastq-mcf (Aronesty, 2011). ERCC spike-in control reads were removed using Bowtie (Langmead and Salzberg, 2012) and quantification of expression was checked to demonstrate quality of the library preparation. Filtered reads were aligned to the draft reference genomes for *T. hamatum* GD12 (Studholme et al., 2013) and unmapped reads were then aligned to the *S. sclerotiorum* strain 1980 (Amselem et al., 2011) using Tophat2 with standard settings (Trapnell et al., 2009).

Statistical analysis of differential expression was carried out using three separate software packages. HTSeq (Anders et al., 2014) was used for quantification of reads at each predicted gene location. These counts were then analysed with the DESeq2 (Love et al., 2014) and edgeR (Robinson et al., 2010) programmes using a significance threshold of an adjusted p value of < 0.05, representing a 5% false discovery rate, corrected for multiple testing using the Benjamini Hochberg method (Benjamini and Hochberg, 1995). Cufflinks (Trapnell et al., 2010) was also used to quantify transcripts, followed by differential expression analysis using CuffDiff2 (Trapnell et al., 2013), using the same significance parameters as above.

Clustering of genes with similar normalised read counts, as determined by DESeq2, was carried out using the clustergram function in MATLAB. Functional predictions of genes were made using results from BLASTX (Altschul et al., 1990) alongside analysis with the Blast2GO software package (Conesa et al., 2005). Predictions of protein location were made using WoLF PSORT (Horton et al., 2007) for all translated gene predictions starting with a methionine. Identification of potential secondary metabolite biosynthetic pathways within the draft reference genome was achieved using antiSMASH version 2.0 (Blin et al., 2013) predictions.

Availability of supporting data Sequencing data supporting the results of this study can be found in the GEO repository under accession number GSE67909.

Author Contributions

SS carried out all analysis of RNA sequencing data, supported by DS, and wrote the manuscript. MG and KL developed the experimental design. KL and RW undertook the experimental assay. KL extracted the RNA. KP and KM carried out the library preparation and sequencing. MZ carried out the qPCR validations. CT and MG advised throughout the investigation and, along with DS and KL, aided in the manuscript preparation.

Acknowledgments

We wish to thank Mark Rowe (Perkin Elmer), Heather Musk and Gawain Bennett (TGAC) for facilitating the sequencing library preparation and Bev Harris for initial contributions to looking at antagonism responses. The authors declare no conflict of

interest. *Funding:* This work was supported by a Biotechnology and Biological Sciences Research Council grant (BB/I014691/1) to MG and CT.

Supplementary Figures

Supplementary Figure 1. Overview of experimental microcosms.

Supplementary Figure 2. Comparison of differential expression calling between DESeq2 and edgeR programmes.

Supplementary Figure 3. RT-PCR validation of selected gene expression patterns between GD12 only microcosms and mixed species microcosms

Supplementary Tables

Supplementary Table 1. Summary of read statistics from each treatment replicate at each of the 6 time points

Supplementary Table 2. Primer combinations used for qPCR validation of selected transcripts from GD12 only and mixed species microcosms (MSMs). SSCRP – small secreted cysteine rich proteins.

Supplementary Table 3. SSCRPs (secreted small cysteine rich proteins) up-regulated in GD12 only microcosms.

Supplementary Table 4. SSCRPs up-regulated in mixed species microcosms.

Supplementary Table 5. Potential secondary metabolite producing gene clusters identified by antiSMASH.

Table legends

Table 1. Number of differentially expressed GD12 genes identified by DESeq2 at each time point in GD12 only or GD12 + *S. sclerotiorum* treatments and genes which are uniquely present to the treatment (normalised count = 0, FDR 5%, BH correction).

Table 2. Number of differentially expressed GD12 genes at concurrent time points during a) GD12 only PGP and b) mixed species antagonism identified by DESeq2 at each time point. Based upon a 5% false discovery rate, using a BH multiple comparison correction.

Table 3. Genes significantly up-regulated in either GD12 only or mixed species microcosms at every time point.

Figure Legends

Figure 1. Summary overview of individual lettuce plants showing plant growth promotion and biocontrol of *Sclerotinia sclerotiorum* by *Trichoderma hamatum* GD12 through the 15 day experimental period. (A) Photographs of wells containing a single representative lettuce plant growing in peat treated with *Trichoderma hamatum* GD12 only, *Sclerotinia sclerotiorum* only, *Trichoderma hamatum* GD12 and *Sclerotinia sclerotiorum* or with no treatment (Control). The diameter of each well is 39 mm. A full overview is presented in Supplementary Figure 1. (B) Quantification of plant growth

promotion by *T. hamatum* GD12 in the presence of *S. sclerotiorum*. Dry weight of lettuce shoots and roots were measured after 14 days of growth in peat supplemented with the treatments presented in A. Each bar isthe mean dry weight per plant of three experimental replicates each containing (25) replicates (plants) ± standard errors. Bars with different letters are significantly different to one another at $p < 0.05$.

Figure 2. Summary of mRNA-seq read distributions across the time course. Proportion of RNA reads at 1 d, 2 d, 4 d, 7 d, 10 d and 15 d aligned to each reference genome. Blue: proportion of reads aligned to the *T. hamatum* GD12 reference. Orange: proportion of unmapped reads then aligned to the *S. sclerotiorum* 1980 reference. Pink: proportion of unmapped reads.

Figure 3. Classifying differentially expressed genes using three differential expression analysis software packages. Number of differentially expressed genes identified by CuffDiff2 (blue), DESeq2 (red) and edgeR (green) at each time point. The top panel shows the number of genes with significantly higher expression in the *T. hamatum* GD12 only microcosms whereas the bottom panel shows the number of genes with significantly higher expression in the mixed species microcosms.

Figure 4. Treatment specific gene dynamics over the experimental time course. Number of genes expressed (normalised count ≥ 25) in either the *T. hamatum* GD12 only microcosms (blue) or the mixed species microcosms (red) at each time point as determined by DESeq2.

Figure 5. Comparing transcriptional dynamics of *T. hamatum* GD12 peat microcosms with mixed peat microcosms of *T. hamatum* GD12 and *S. sclerotiorum*. (A) The transcriptional response in GD12 only microcosm (blue) peak at 2 d and 15 d, and the

transcriptional response in the mixed species microcosms (red) peak at 2 d and 10 d. The dashed lines show that the number of differentially expressed genes with counts solely in one sample, with no expression in the other, and is very few for both GD12 only and mixed species. (B) The transcriptional response of *T. hamatum* GD12 is highly dynamic. Genes clustered by normalised counts using clustergram in MATLAB. Red shows a high level of expression whereas green shows a low level of expression. (C) Percentage of DEGs associated with the GO terms transport and oxidation-reduction is higher in mixed species microcosms at initial time points up to 4 d.

Figure 6. Hierarchal clustering reveals temporal deployment of specific sets of gene in mixed species microcosms. Clustering of genes based on normalised counts of reads using clustergram in MATLAB. (A) 18 genes with higher expression in the mixed species microcosms on 1 d, 2 d and 4 d. (B) 103 genes with higher expression in the MSMs on 4 d, 7 d, 10 d and 15 d.

Figure 7. Predicted locations of Small Secreted Cysteine Rich Peptides. (A) Percent of location, as predicted by WoLF PSORT, of all predicted SSCRPs (green; Studholme *et al.* 2013). Those up-regulated in mixed species microcosms are coloured blue and those up-regulated in GD12 only microcosms, red. Predicted proteins lacking a methionine start could not be processed by WoLF PSORT, and are therefore not present. (B) Number of SSCRPs genes up-regulated at each predicted location at each time point in the mixed species microcosms.

Figure 8. Expression trends for predicted genes with significant differences in expression between GD12 only and mixed species microcosms on the following contigs: (A) ANCB01004730.1, (B) ANCB01008794.1, (C) ANCB010770.1, (D) ANCB01011924.1, (E) ANCB01005027.1, (F) ANCB01007958.1, (G)

ANCB01007961.1, and (H) ANCB010115.1. Positive-fold change shows higher expression in the mixed species microcosms whereas a negative-fold change shows higher expression in the GD12 only microcosms. Arrangement of gene cluster on contig as predicted by antiSMASH is shown below. Annotation of cluster was by antiSMASH and gene prediction by Blast2GO.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J Mol Biol **215,** 403-410.

Amselem, J., **Cuomo, C.A., van Kan, J.A.L., Viaud, M., Benito, E.P., Couloux, A., Coutinho, P.M., de Vries, R.P., et al.** (2011) Genomic analysis of the nectrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea.* PLoS Genet. **18,** 7(8):e1002230.

Anders, S., Pyl, P.T., and Huber, W. (2014). HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics **31,** 166-169

Anke, H., Kinn, J., Bergquist, K.-E., and Sterner, O. (1991). Production of siderophores by strains of the genus *Trichoderma*: Isolate and characterization of the new lipophilic coprogen derivative, palmitoylcoprogen. Biol. Met. **4,** 176-180.

Aronesty, E. (2011). ea-utils: "Command-line tools for processing biological sequencing data; http://code.google.com/p/ea-utils.

Atanasova, L., Crom, S.L., Gruber, S., Coulpier, F., Seidl-Seiboth, V., Kubicek, C., and Druzhinina, I. (2013). Comparative transcriptomics reveals different strategies of *Trichoderma* mycoparasitism. BMC Genomics **14,** 121.

Bae, H., Sicher, R.C., Kim, M.S., Kim, S.-H., Strem, M.D., Melnick, R.L., and Bailey, B.A. (2009). The beneficial endophyte *Trichoderma hamatum* isolate DIS 219b promotes growth and delays the onset of the drought response in Theobroma cacao. J. Exp. Bot. **60,** 3279-3295.

Baker, S.E., Perrone, G., Richardson, N.M., Gallo, A., and Kubicek, C.P. (2012). Phylogenomic analysis of polyketide synthase-encoding genes in Trichoderma. Microbiology **158,** 147-154.

Benítez, T., Rincón, A.M., Limón, M.C. and Codón, A.C. (2004) Biocontrol mechanisms of *Trichoderma* strains. International Microbiology **7,** 249-260.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. J. R. Statist. Soc. B **57,** 289-300.

Blin, K., Medema, M.H., Kazempour, D., Fischbach, M.A., Breitling, R., Takano, E., and Weber, T. (2013). antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. Nucleic Acids Res. **41,** W204-W212.

Brotman, Y., Landau, U., Cuadros-Inostroza, Á., Takayuki, T., Fernie, A.R., Chet, I., Viterbo, A., and Willmitzer, L. (2013). *Trichoderma*-Plant Root Colonization: Escaping Early Plant Defense Responses and Activation of the Antioxidant Machinery for Saline Stress Tolerance. PLoS Pathog. **9,** e1003221.

Cardoza, R.E., Vizcaíno, J.A., Hermosa, M.R., Sousa, S., González, F.J., Llobell, A., Monte, E., and Gutiérrez, S. (2006). Cloning and characterization of the *erg1* gene of *Trichoderma harzianum*: Effect of the *erg1* silencing on ergosterol biosynthesis and resistance to terbinafine. Fungal Genet. Biol. **43,** 164-178.

Carpenter, M.A., Stewart, A., and Ridgway, H.J. (2005). Identification of novel *Trichoderma hamatum* genes expressed during mycoparasitism using subtractive hybridisation. FEMS Microbiol. Lett. **251,** 105-112.

Carvalho, F.P. (2006) Agriculture, pesticides, food security and food safety. Environmental Science and Policy **9,** 685-692.

Chang, Y.-C., Chang, Y.-C., and Baker, R. (1986). Increased growth of plants in the prescence of the biological control agent *Trichoderma harzianum*. Plant Dis. **70,** 145-148.

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics **21,** 3674-3676.

Donoso, E.P., Bustamante, R.O., Carú, M., and Niemeyer, H.M. (2008). Water Deficit as a Driver of the Mutualistic Relationship between the Fungus *Trichoderma harzianum* and Two Wheat Genotypes. Applied and Environ. Microbiol. **74,** 1412-1417.

Frewer, L., Lassen, J., Kettlitz, B., Scholderer, J., Beekman, V., and Berdal, K.G. (2004). Societal aspects of genetically modified foods. Food Chem. Toxicol. **42,** 1181-1193.

Hermosa, R., Viterbo, A., Chet, I., and Monte, E. (2012). Plant-beneficial effects of *Trichoderma* and of its genes. Microbiology **158,** 17-25.

Hillier, J., Hawes, C., Squire, G., Hilton, A., Wale, S., and Smith, P. (2009). The carbon footprints of food crop production. International Journal of Agricultural Sustainability **7,** 107-118.

Horton, P., Park, K.-J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., and Nakai, K. (2007). WoLF PSORT: protein localization predictor. Nucleic Acids Res. **35,** W585-W587.

Howell, C.R. (2003). Mechanisms Employed by *Trichoderma* Species in the Biological Control of Plant Diseases: The History and Evolution of Current Concepts. Plant Disease **87,** 4-10.

Hoyos-Carvajal, L., Orduz, S., and Bissett, J. (2009). Growth stimulation in bean (*Phaseolus vulgaris* L.) by *Trichoderma*. Biological Control **51,** 409-416.

Hu, B., Li, S., Zhang, X., and Zheng, X. (2014). HSCARG, a novel regulator of H2A ubiquitination by downregulating PRC1 ubiquitin E3 ligase activity, is essential for cell proliferation. Nucleic Acids Res. **42,** 5582-5593.

Huesing, J.E., Andres, D., Braverman, M.P., Burns, A., Felsot, A.S., Harrigan, G.G., Hellmich, R.L., Reynolds, A., Shelton, A.M., Jansen van Rijssen, W., and Morris, E.J. (2016). Global Adoption of Genetically Modified (GM) Crops: Challenges for the Public Sector. J. Agric. Food Chem. **64,** 394-402.

Kimura, M., Kaneko, I., Komiyama, M., Takatsuki, A., Koshino, H., Yoneyama, K., and Yamaguchi, I. (1998). Trichothecene 3-O-acetyltransferase protects both the producing organism and transformed yeast from related mycotoxins. Cloning and characterization of Tri101. The Journal of Biological Chemistry **273,** 1654-1661.

Lamdan, N.-L., Shalaby, S., Ziv, T., Kenerley, C.M., and Horwitz, B.A. (2015). Secretome of the biocontrol fungus *Trichoderma virens* co-cultured with maize roots: role in induced systemic resistance. Mol. Cell. Proteomics **14,** 1054-1063.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods **9,** 357-359.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Research **15,** 550.

Magnusson, M.K., and Hursti, U.K.K. (2002). Consumer attitudes towards genetically modified foods. Appetite **39,** 9-24.

Mastouri, F., Björkman, T., and Harman, G.E. (2010). Seed Treatment with *Trichoderma harzianum* Alleviates Biotic, Abiotic, and Physiological Stresses in Germinating Seeds and Seedlings. Phytopathology **100,** 1213-1221.

Matarese, F., Sarrocco, S., Gruber, S., Seidl-Seiboth, V., and Vannacci, G. (2012). Biocontrol of Fusarium head blight: interactions between Trichoderma and mycotoxigenic Fusarium. Microbiology **158,** 98-106.

Meyer, J.-M. (2000). Pyoverdines: pigments, siderophores and potential taxonomic markers of fluorescent *Pseudomonas* species. Archives of Microbiology **174,** 135-142.

Moran-Diez, M.E., Trushina, N., Lamdan, N.L., Rosenfelder, L., Murkherjee, P.K., Kenerley, C.M., and Horwitz, B.A. (2015). Host-specific transcriptomic pattern of *Trichoderma virens* during interaction with maize or tomato roots. BMC Genomics **16,** 8.

Mukherjee, P.K., Horwitz, B.A., and Kenerley, C.M. (2012). Secondary metabolism in Trichoderma--a genomic perspective. Microbiology **158,** 35-45.

Nookaew, I., Papini, M., Pornputtapong, N., Scalcinati, G., Fagerberg, L., Uhlén, M., and Nielsen, J. (2012). A comprehensive comparison of RNA-Seq-based transcriptome analysis from reads to differential gene expression and cross-comparison with microarrays: a case study in *Saccharomyces cerevisiae*. Nucleic Acids Res. **40,** 10084-10097.

Oerke, E., and Dehne, H. (2004). Safeguarding production—losses in major crops and the role of crop protection. Crop Protection **23**, 275-285

Omann, M.R., Lehner, S., Escobar Rodriguez, C., Brunner, K., and Zeilinger, S. (2012). The seven-transmembrane receptor Gpr1 governs processes relevant for the antagonistic interaction of Trichoderma atroviride with its host. Microbiology **158,** 107-118.

Pharmaceutica, B. (2002). Aerothricins: a new class of β-glucan inhibitors. Expert Opinion on Therapeutic Patents **12,** 315-318.

Renshaw, J.C., Robson, G.D., Trinci, A., P.J., Wiebe, M., G., Livens, F., R., Collison, D., and Taylor, R.J. (2002). Fungal siderophores: structures, functions and applications. Mycol. Res. **106,** 1123-1142.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **26,** 139-140.

Ryder, L.S., Harris, B.D., Soanes, D.M., Kershaw, M.J., Talbot, N.J., and Thornton, C.R. (2012). Saprotrophic competitiveness and biocontrol fitness of a genetically modified strain of the plant-growth-promoting fungus *Trichoderma hamatum* GD12. Microbiology **158,** 84-97.

Samolski, I., de Luis, A., Vizcaíno, J.A., Monte, E., and Suárez, M.B. (2009). Gene expression analysis of the biocontrol fungus *Trichoderma harzianum* in the presence of tomato plants, chitin, or glucose using a high-density oligonucleotide microarray. BMC Microbiol. **9,** 217-217.

Savary, S., Ficke, A., Aubertot, J.-N., and Hollier, C. (2012). Crop losses due to diseases and their implications for global food production losses and food security. Food Security **4,** 519-537.

Schlömann, M., Ngai, K.L., Ornston, L.N., and Knackmuss, H.J. (1993). Dienelactone hydrolase from *Pseudomonas cepacia*. Journal of Bacteriology **175,** 2994-3001.

Schmidt, E., and Knackmuss, H.J. (1980). Chemical structure and biodegradability of halogenated aromatic compounds. Conversion of chlorinated muconic acids into maleoylacetic acid. Biochemical Journal **192,** 339-347.

Steindorff, A., Ramada, M.H., Coelho, A.S., Miller, R.N., Pappas, G., Ulhoa, C., and Noronha, E. (2014). Identification of mycoparasitism-related genes against the phytopathogen *Sclerotinia sclerotiorum* through transcriptome and expression profile analysis in *Trichoderma harzianum*. BMC Genomics **15,** 204.

Steyaert, J.M., Stewart, A., Jaspers, M.V., Carpenter, M., and Ridgway, H.J. (2004). Co-expression of two genes, a chitinase (*chit42*) and proteinase (*prb1*), implicated in mycoparasitism by *Trichoderma hamatum*. Mycologia **96,** 1245-1252.

Studholme, D.J., Harris, B., Le Cocq, K., Winsbury, R., Perera, V., Ryder, L., Beale, M., Ward, J., Thornton, C.R., and Grant, M. (2013). Investigating the beneficial traits of *Trichoderma hamatum* GD12 for sustainable agriculture – insights from genomics. Frontiers in Plant Science **4**, 258.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics **25,** 1105-1111.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat. Biotechnol. **31,** 46-53.

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

Verma, V.C., Singh, S.K., and Prakash, S. (2011). Bio-control and plant growth promotion potential of siderophore producing endophytic Streptomyces from Azadirachta indica A. Juss. J. Basic Microbiol. **51,** 550-556.

Vos, C.M.F., De Cremer, K., Cammue, B.P.A., and De Coninck, B. (2014). The toolbox of *Trichoderma* spp. in the biocontrol of *Botrytis cinerea* disease. Mol. Plant Pathol**. 16,** 400-412.

Yedidia, I., Srivastva, A., Kapulnik, Y., and Chet, I. (2001). Effect of *Trichoderma harzianum* on microelement concentrations and increased growth of cucumber plants. Plant and Soil **235,** 235-242.

Yu, X., Ai, C., Xin, L., and Zhou, G. (2011). The siderophore-producing bacterium, Bacillus subtilis CAS15, has a biocontrol effect on Fusarium wilt and promotes the growth of pepper. Eur. J. Soil Biol. **47,** 138-145.

Zhang, Z.H., Jhaveri, D.J., Marshall, V.M., Bauer, D.C., Edson, J., Narayanan, R.K., Robinson, G.J., Lundberg, A.E., Bartlett, P.F., Wray, N.R., and Zhao, Q.-Y. (2014). A Comparative Study of Techniques for Differential Expression Analysis on RNA-Seq Data. PLoS ONE **9,** e103207.

Zhao, Y., Zhang, J., Li, H., Li, Y., Ren, J., Luo, M., and Zheng, X. (2008). An NADPH Sensor Protein (HSCARG) Down-regulates Nitric Oxide Synthesis by Association with Argininosuccinate Synthetase and Is Essential for Epithelial Cell Viability. Journal of Biological Chemistry **283,** 11004-11013.

Table 1. Number of differentially expressed GD12 genes identified by DESeq2 at each time point in GD12 only or GD12 + *S. sclerotiorum* treatments and genes which are uniquely present to the treatment (normalised count = 0, FDR 5%, BH correction).

Time Point	Number of Genes Up- Regulated In Peat In GD12 only	Number of Genes Up- Regulated In Peat In GD12 only where lower expression is 0	Number of Genes Up- Regulated In Peat In Mix	Number of Genes Up Regulated In Peat In Mix where lower expression is 0
Day 1	454	1	238	0
Day 2	1905	2	891	7
Day 4	1378	0	442	3
Day 7	873	2	993	10
Day 10	669	2	2351	42
Day 15	1952	0	1808	58

Table 2. Number of differentially expressed GD12 genes at concurrent time points during a) GD12 only PGP and b) mixed species antagonism identified by DESeq2 at each time point. Based upon a 5% false discovery rate, using a BH multiple comparison correction.

a) GD12 only	**1 d**	**1 & 2 d**	**1, 2 & 4 d**	**1, 2, 4 & 7 d**	**1, 2, 4, 7 & 10 d**	**1, 2, 4, 7, 10 & 15 d**
	454	51	21	11	9	9
		2 d	**2 & 4 d**	**2, 4 & 7 d**	**2, 4, 7 & 10 d**	**2, 4, 7, 10 & 15 d**
		1905	293	47	23	22
			4 d	**4 & 7 d**	**4, 7 & 10 d**	**4, 7, 10 & 15 d**
			1378	306	149	135
				7 d	**7 & 10 d**	**7, 10 & 15 d**
				873	399	370
					10 d	**10 & 15 d**
					669	562
						15 d
						1952
b) Mixed Species	**1 d**	**1 & 2 d**	**1, 2 & 4 d**	**1, 2, 4 & 7 d**	**1, 2, 4, 7 & 10 d**	**1, 2, 4, 7, 10 & 15 d**
	238	87	48	29	26	22
		2 d	**2 & 4 d**	**2, 4 & 7 d**	**2, 4, 7 & 10 d**	**2, 4, 7, 10 & 15 d**
		891	187	104	98	88
			4 d	**4 & 7 d**	**4, 7 & 10 d**	**4, 7, 10 & 15 d**
			442	220	210	176
				7 d	**7 & 10 d**	**7, 10 & 15 d**
				993	804	613
					10 d	**10 & 15 d**
					2351	1258
						15 d
						1808

Table 3. Genes significantly up-regulated in either GD12 only or mixed species microcosms at every time point.

Up Regulated During GD12 Only			Up Regulated During Mixed Species		
Gene Name (ANCB0XXXX)	**Log_2 Fold Change at Day One (to 2 d.p.)**	**BLASTX Hit**	**Gene Name (ANCBXXXX)**	**Log_2 Fold Change at Day One (to 2 d.p.)**	**BLASTX Hit**
1006842.1:599-2594	5.19	NAD dependent epimerase dehydratase	1009877.1:936-3483	6.19	
1003985.1:771-3990	1.98	Malic acid transport protein	1007557.1:3442-5242	5.65	
1002344.1:7484-8547	1.85	ubie coq5 methyltransferase	1005725.1:65-1614	4.96	ABC multidrug
1010824.1:460-1942	1.57		1011666.1:16-764	4.62	
1002225.1:1802-2938	1.47	GNAT family protein	1010015.1:968-2443	4.48	Pyoverdine dityrosine biosynthesis
1006475.1:142-1395	1.20	Methyltransferase-like protein	1011665.1:29-942	4.43	
1013852.1:19-1423	1.19	Tam domain	1009227.1:295-2134	3.7	Dienelactone hydrolase family protein
1002254.1:150-2104	1.18		1013178.1:264-4974	3.67	Transferase family protein
1003305.1:15-928	0.92	Potassium uptake protein	1013143.1:5756-7036	3.56	Methyltransferase domain-containing protein
			1004667.1:16-788	3.55	
			1001608.1:249-1661	3.47	Mitochondrial phosphate carrier protein
			1014508.1:600-2990	3.39	Short chain dehydrogenase reductase
			1013443.1:1767-2642	3.33	Fluconazole resistance protein 1
			1002135.1:4185-5616	3.24	
			1010016.1:125-2333	2.92	MFS transporter
			1009878.1:815-3132	2.91	Family regulatory protein
			1000041.1:739-1491	2.50	FAD linked oxidase
			1010706.1:2182-4056	2.40	GNAT family
			1009777.1:2530-3653	2.08	Dienelactone hydrolase family protein
			1005944.1:2672-3858	1.86	FAD dependent oxidoreductase
			1005794.1:2226-3132	1.83	Aryl-alcohol dehydrogenase
			1003990.1:109-1714	1.21	Major facilitator superfamily transporter

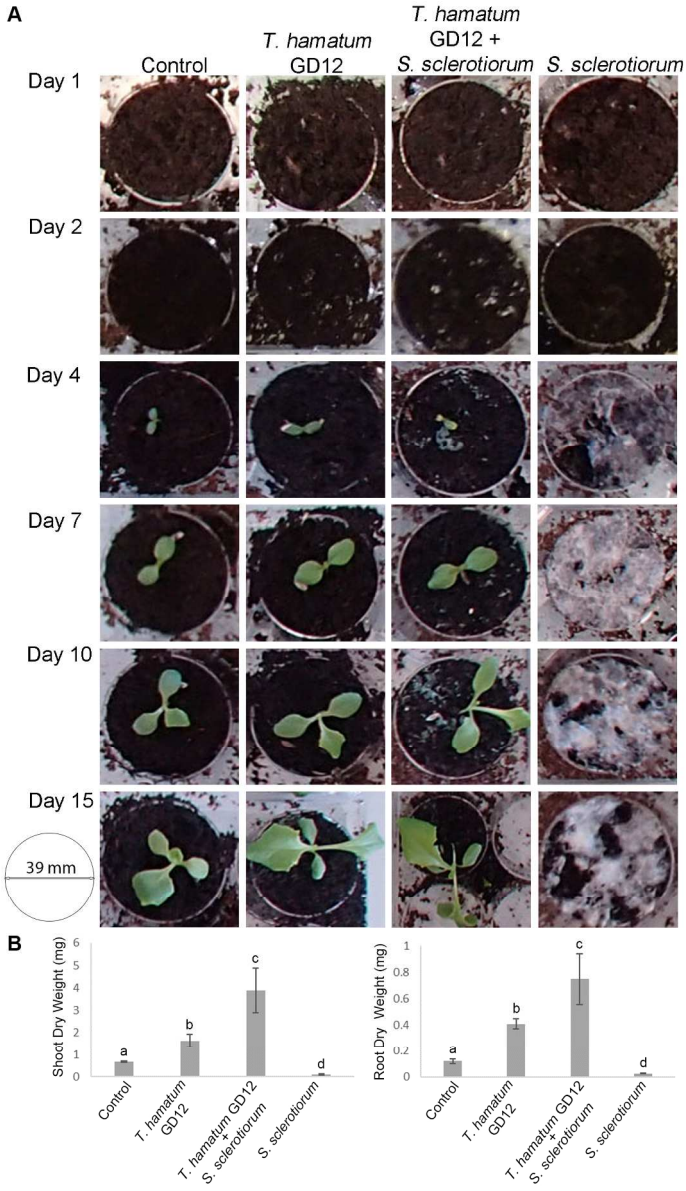


Figure 1. Summary overview of individual lettuce plants showing plant growth promotion and biocontrol of Sclerotinia sclerotiorum by Trichoderma hamatum GD12 through the 15 day experimental period. (A) Photographs of wells containing a single representative lettuce plant growing in peat treated with Trichoderma hamatum GD12 only, Sclerotinia sclerotiorum only, Trichoderma hamatum GD12 and Sclerotinia sclerotiorum or with no treatment (Control). The diameter of each well is 39 mm. A full overview is presented in Supplementary Figure 1. (B) Quantification of plant growth promotion by T. hamatum GD12 in the presence of S. sclerotiorum. Dry weight of lettuce shoots and roots were measured after 14 days of growth in peat supplemented with the treatments presented in A. Each bar is the mean dry weight per plant of three experimental replicates each containing (25) replicates (plants) ± standard errorse. Bars with different letters are significantly different to one another at $p < 0.05$.
144x245mm (300 x 300 DPI)

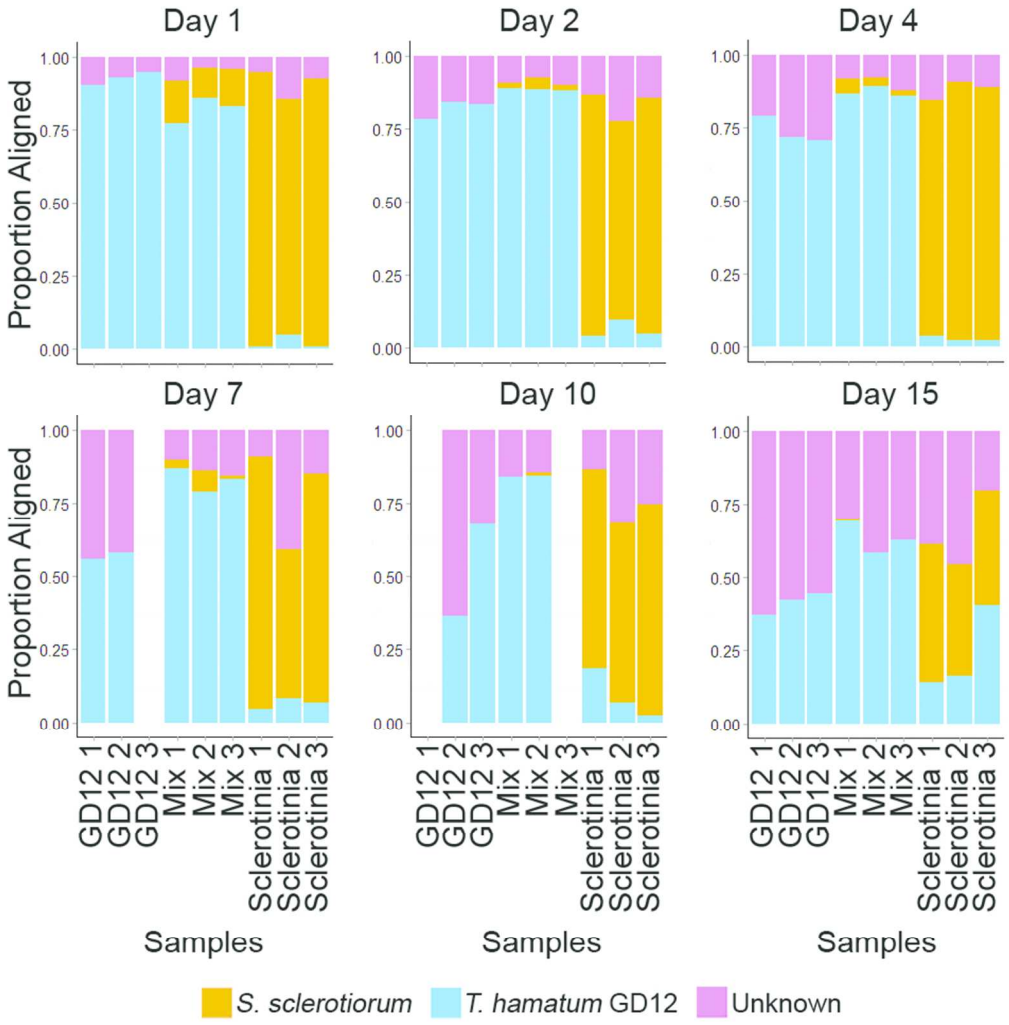


Figure 2. Summary of mRNA-seq read distributions across the time course. Proportion of RNA reads at 1 d, 2 d, 4 d, 7 d, 10 d and 15 d aligned to each reference genome. Blue: proportion of reads aligned to the T. hamatum GD12 reference. Orange: proportion of unmapped reads then aligned to the S. sclerotiorum 1980 reference. Pink: proportion of unmapped reads.
80x83mm (300 x 300 DPI)

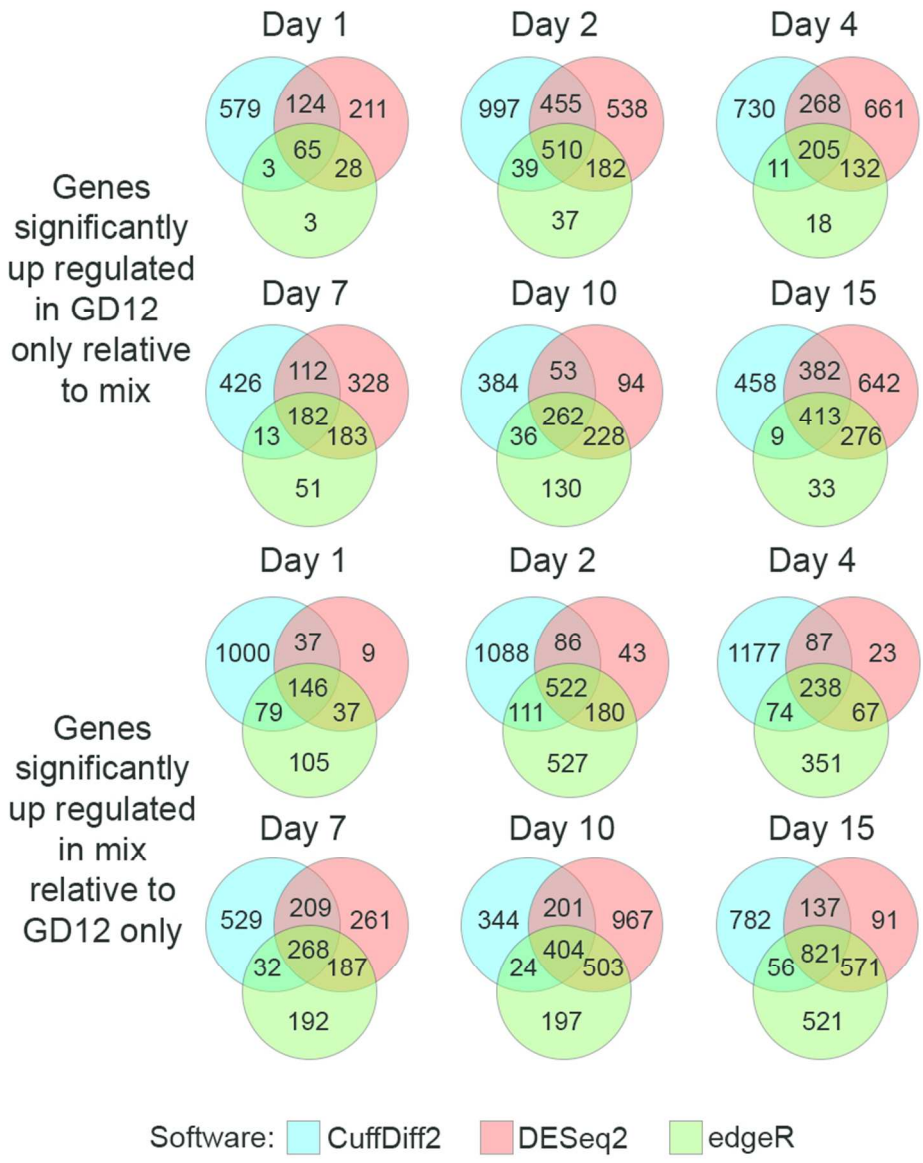


Figure 3. Classifying differentially expressed genes using three differential expression analysis software packages. Number of differentially expressed genes identified by CuffDiff2 (blue), DESeq2 (red) and edgeR (green) at each time point. The top panel shows the number of genes with significantly higher expression in the T. hamatum GD12 only microcosms whereas the bottom panel shows the number of genes with significantly higher expression in the mixed species microcosms.
80x99mm (300 x 300 DPI)

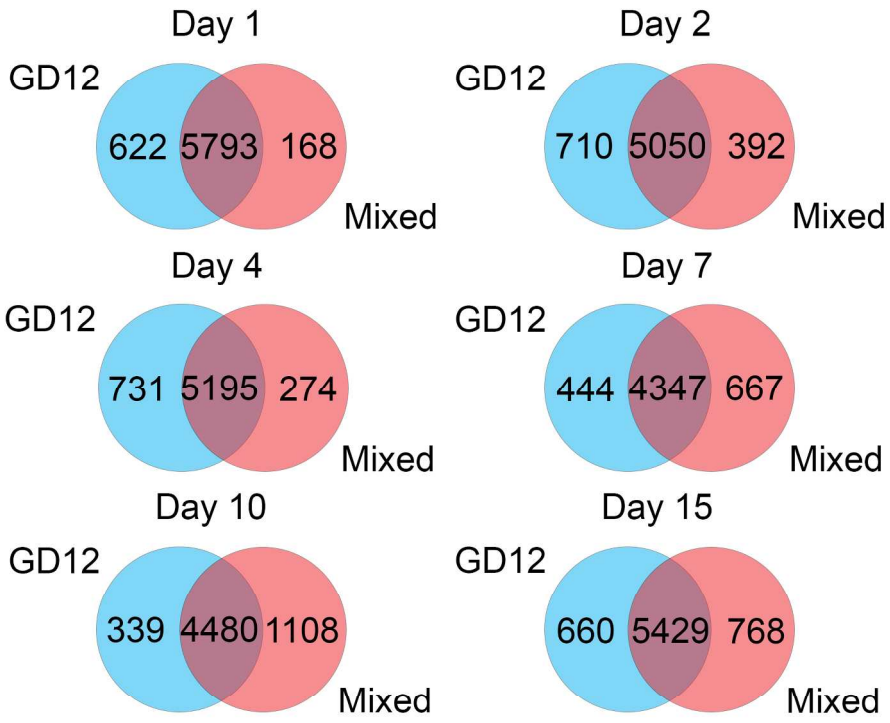


Figure 4. Treatment specific gene dynamics over the experimental time course. Number of genes expressed (normalised count ≥ 25) in either the T. hamatum GD12 only microcosms (blue) or the mixed species microcosms (red) at each time point as determined by DESeq2.

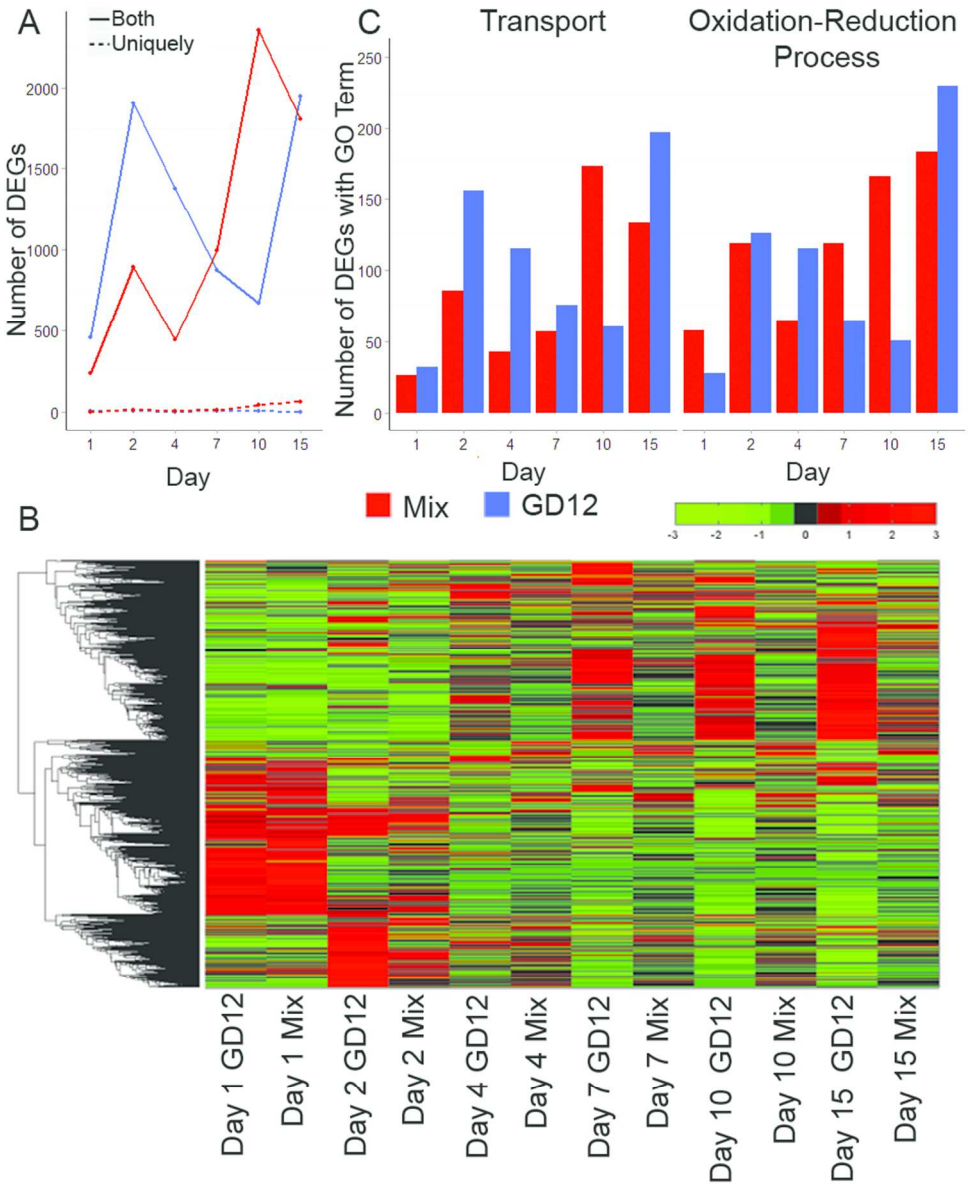


Figure 5. Comparing transcriptional dynamics of T. hamatum GD12 peat microcosms with mixed peat microcosms of T. hamatum GD12 and S. sclerotiorum. (A) The transcriptional response in GD12 only microcosm (blue) peak at 2 d and 15 d, and the transcriptional response in the mixed species microcosms (red) peak at 2 d and 10 d. The dashed lines show that the number of differentially expressed genes with counts solely in one sample, with no expression in the other, and is very few for both GD12 only and mixed species. (B) The transcriptional response of T. hamatum GD12 is highly dynamic. Genes clustered by normalised counts using clustergram in MATLAB. Red shows a high level of expression whereas green shows a low level of expression. (C) Percentage of DEGs associated with the GO terms transport and oxidation-reduction is higher in mixed species microcosms at initial time points up to 4 d.
80x99mm (300 x 300 DPI)

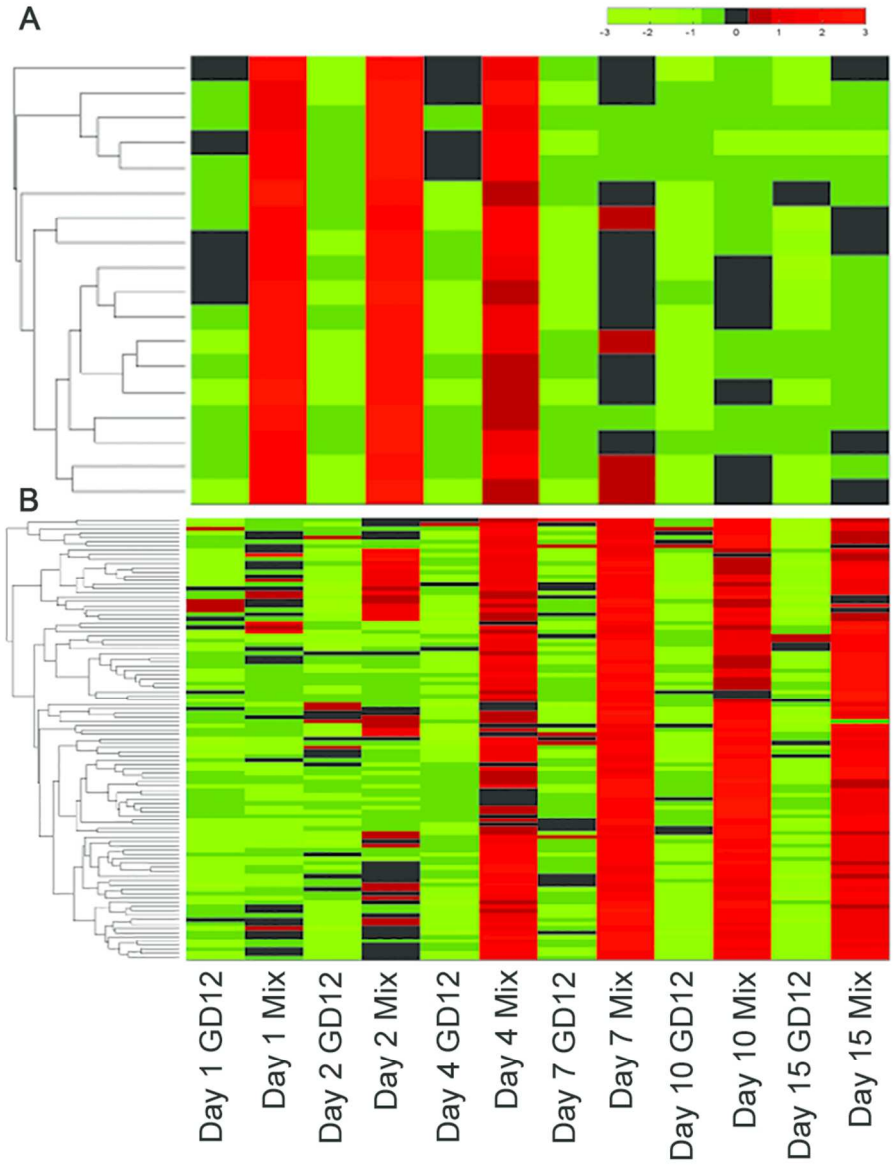
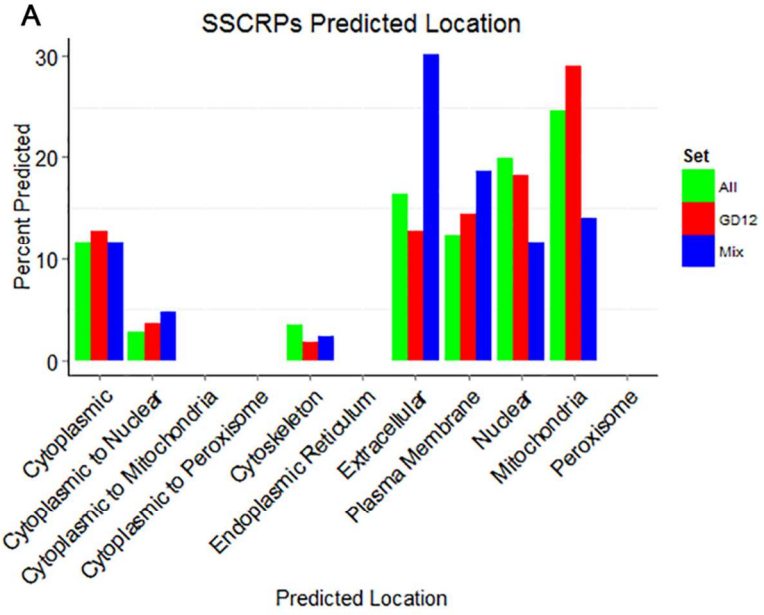


Figure 6. Hierarchal clustering reveals temporal deployment of specific sets of gene in mixed species microcosms. Clustering of genes based on normalised counts of reads using clustergram in MATLAB. (A) 18 genes with higher expression in the mixed species microcosms on 1 d, 2 d and 4 d. (B) 103 genes with higher expression in the MSMs on 4 d, 7 d, 10 d and 15 d.
80x107mm (300 x 300 DPI)

A



B

Number of Genes Up Regulated in Mix at Each Time Point

Day	Cytoplasmic	Cytoplasmic to Nuclear	Cytoplasmic to Mitochondria	Cytoplasmic to Peroxisome	Cytoskeleton	Endoplasmic Reticulum	Extracellular	Plasma Membrane	Nuclear	Mitochondria	Peroxisome
1	1	1					1				
2		2					7		1	1	
4	1	2					2	2	1		
7	3	1					3	4	2	2	
10	3	1					6	6	3	3	
15	2	1			1		8	6		3	

Figure 7. Predicted locations of Small Secreted Cysteine Rich Peptides. (A) Percent of location, as predicted by WoLF PSORT, of all predicted SSCRPs (green; Studholme et al. 2013). Those up-regulated in mixed species microcosms are coloured blue and those up-regulated in GD12 only microcosms, red. Predicted proteins lacking a methionine start could not be processed by WoLF PSORT, and are therefore not present. (B) Number of SSCRPs genes up-regulated at each predicted location at each time point in the mixed species microcosms.
80x122mm (300 x 300 DPI)

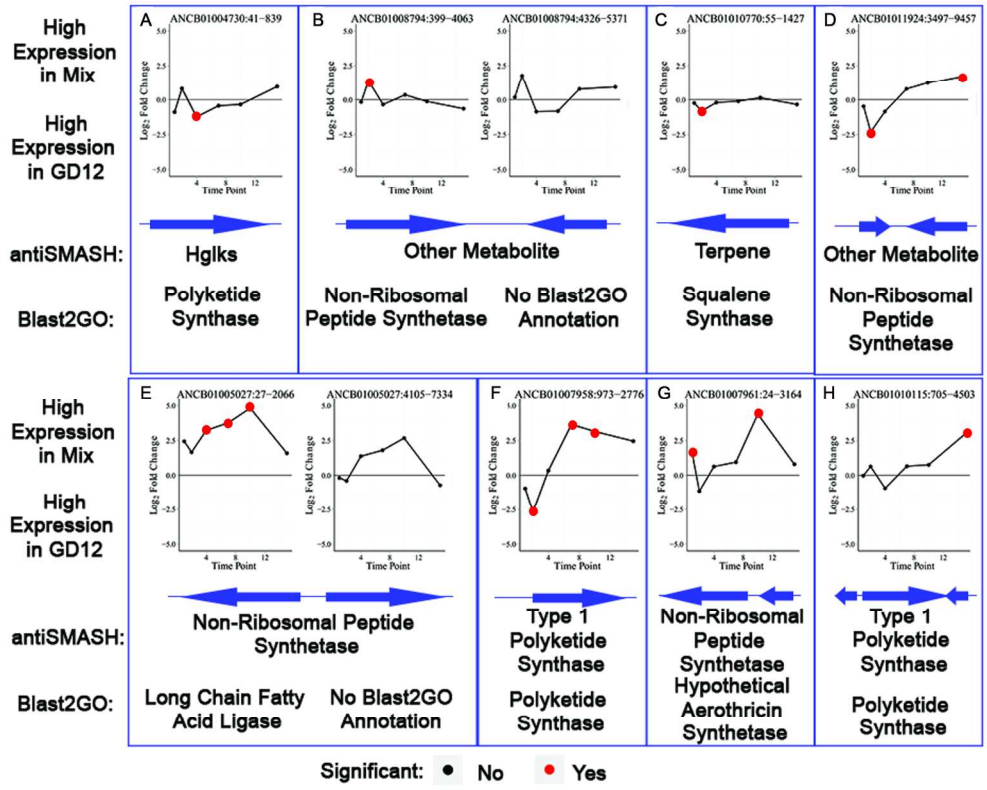


Figure 8. Expression trends for predicted genes with significant differences in expression between GD12 only and mixed species microcosms on the following contigs: (A) ANCB01004730.1, (B) ANCB01008794.1, (C) ANCB010770.1, (D) ANCB01011924.1, (E) ANCB01005027.1, (F) ANCB01007958.1, (G) ANCB01007961.1, and (H) ANCB010115.1. Positive-fold change shows higher expression in the mixed species microcosms whereas a negative-fold change shows higher expression in the GD12 only microcosms. Arrangement of gene cluster on contig as predicted by antiSMASH is shown below. Annotation of cluster was by antiSMASH and gene prediction by Blast2GO.
160x135mm (300 x 300 DPI)