

RESEARCH ARTICLE

Open Access



# The genome of the emerging barley pathogen *Ramularia collo-cygni*

Graham R. D. McGrann<sup>1\*</sup>, Ambrose Andongabo<sup>2†</sup>, Elisabet Sjökvist<sup>1,4†</sup>, Urmi Trivedi<sup>5</sup>, Francois Dussart<sup>1</sup>, Maciej Kaczmarek<sup>1,6</sup>, Ashleigh Mackenzie<sup>1</sup>, James M. Fountaine<sup>1,7</sup>, Jeanette M. G. Taylor<sup>1</sup>, Linda J. Paterson<sup>1</sup>, Kalina Gorniak<sup>1</sup>, Fiona Burnett<sup>1</sup>, Kostya Kanyuka<sup>3</sup>, Kim E. Hammond-Kosack<sup>3</sup>, Jason J. Rudd<sup>3</sup>, Mark Blaxter<sup>4,5</sup> and Neil D. Havis<sup>1</sup>

## Abstract

**Background:** *Ramularia collo-cygni* is a newly important, foliar fungal pathogen of barley that causes the disease Ramularia leaf spot. The fungus exhibits a prolonged endophytic growth stage before switching life habit to become an aggressive, necrotrophic pathogen that causes significant losses to green leaf area and hence grain yield and quality.

**Results:** The *R. collo-cygni* genome was sequenced using a combination of Illumina and Roche 454 technologies. The draft assembly of 30.3 Mb contained 11,617 predicted gene models. Our phylogenomic analysis confirmed the classification of this ascomycete fungus within the family Mycosphaerellaceae, order Capnodiales of the class Dothideomycetes. A predicted secretome comprising 1053 proteins included redox-related enzymes and carbohydrate-modifying enzymes and proteases. The relative paucity of plant cell wall degrading enzyme genes may be associated with the stealth pathogenesis characteristic of plant pathogens from the Mycosphaerellaceae. A large number of genes associated with secondary metabolite production, including homologs of toxin biosynthesis genes found in other Dothideomycete plant pathogens, were identified.

**Conclusions:** The genome sequence of *R. collo-cygni* provides a framework for understanding the genetic basis of pathogenesis in this important emerging pathogen. The reduced complement of carbohydrate-degrading enzyme genes is likely to reflect a strategy to avoid detection by host defences during its prolonged asymptomatic growth. Of particular interest will be the analysis of *R. collo-cygni* gene expression during interactions with the host barley, to understand what triggers this fungus to switch from being a benign endophyte to an aggressive necrotroph.

**Keywords:** Ramularia leaf spot, Dothideomycetes, Rubellin toxin, Endophyte, Necrotroph, Whole genome sequencing

## Background

Ramularia leaf spot has emerged as a newly important disease of barley associated with significant grain yield and quality losses across Europe and a number of other temperate regions of the world [1]. The disease was first recognised in 1893 in Italy and the fungal pathogen first described as *Ophiocladium hordei* [2]. Sutton and Waller [3] reclassified this ascomycete fungus to the genus *Ramularia*, within the family Mycosphaerellaceae in the class Dothideomycetes and proposed the species

name *R. collo-cygni* because of the distinctive swan's neck-like shape of the fungal conidiophores (Fig. 1a, b). This systematic position was subsequently supported by phylogenetic analysis of fungal ribosomal DNA sequences [4, 5].

The importance of Ramularia leaf spot as a disease of barley has become more apparent in recent years as reliable methods for isolation and detection of *R. collo-cygni* have become available [6–10]. Ramularia leaf spot is often confused with physiological leaf spotting and leaf spotting diseases, such as net blotch caused by the fungus *Pyrenophora teres*, but is distinguished by characteristic reddish-brown, rectangular lesions visible on both sides of the leaf and surrounded by a chlorotic halo

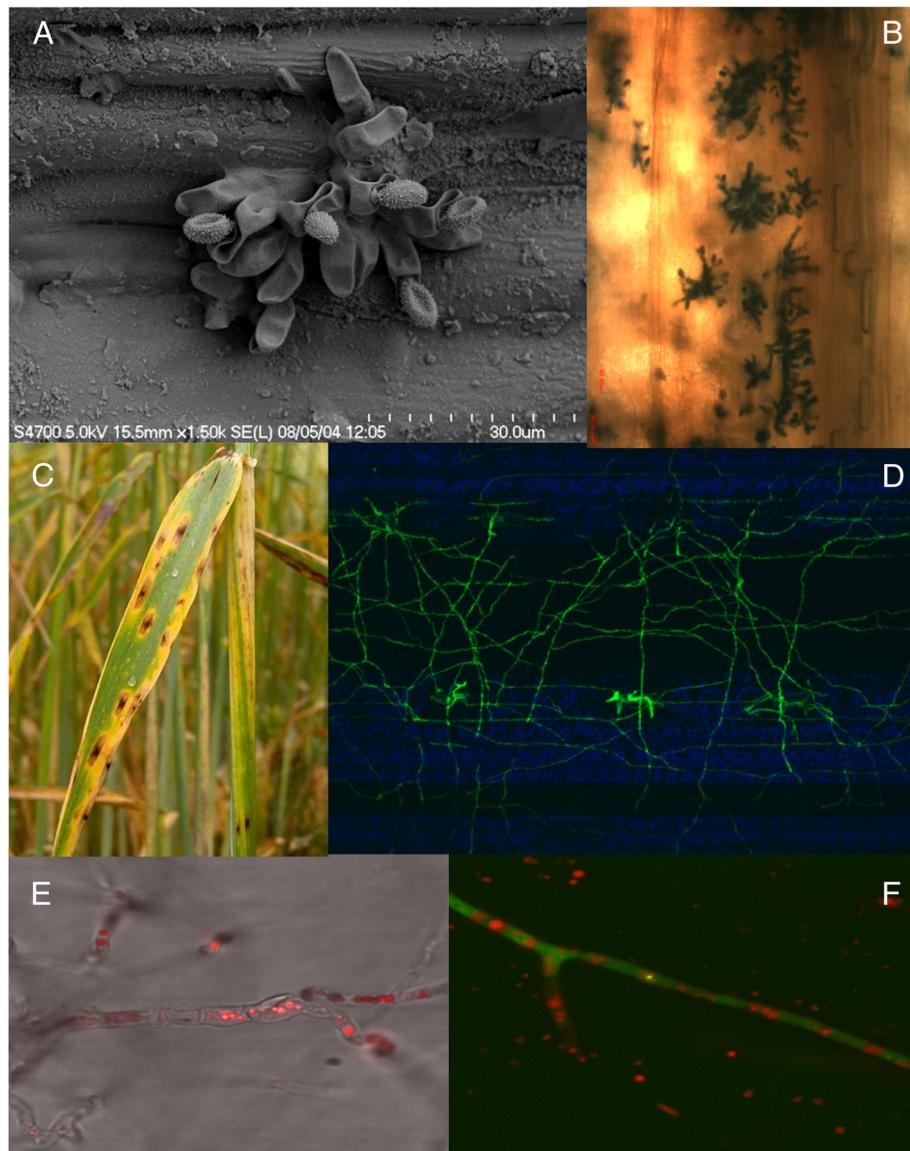
\* Correspondence: graham.mcgrann@sruc.ac.uk

†Equal contributors

<sup>1</sup>Crop Protection Team, Crop and Soil Systems Group, SRUC, West Mains Road, Edinburgh EH9 3JG, UK

Full list of author information is available at the end of the article





**Fig. 1** *Ramularia* leaf spot symptoms and *R. collo-cygni* development on barley. **a** Electron micrograph of the characteristic swan neck-shaped conidiophore of *Ramularia collo-cygni*. **b** Light micrograph of a cluster of *R. collo-cygni* conidiophores, from field infected plants, erupting from spring barley cv. Cocktail leaves at GS80+. **c** Characteristic *Ramularia* leaf spot symptoms on leaves of winter barley cv. Saffron. **d** Spider-web growth of GFP-transformed *R. collo-cygni* isolate Rcc-8B9-GFP [10] on leaf surface of spring barley cv. Optic, 7 days post infection without triggering host immune response. **e** Confocal image of rubellin autofluorescence in *R. collo-cygni* isolate B1, superimposed on bright field transmission image. Note various sizes of vacuole-like vesicles within the collapsed hyphae. **f** Confocal image of a viable hypha of the transgenic isolate Rcc-8B9-GFP carrying rubellin vesicles

(Fig. 1c). Recently reclassified as a major pathogen of its main host plant barley (*Hordeum vulgare*), *R. collo-cygni* has also been reported to colonise wheat, rye, oats, maize and many other grass species including the model *Brachypodium distachyon* [11, 12]. *R. collo-cygni* is transmitted both via spores and through infected seeds [6, 13], and exhibits intercellular, apoplastic colonisation of its host [10, 14, 15]. Disease is typically observed late in the growing season after the crop flowers, suggesting

a link between *Ramularia* leaf spot symptom expression and host developmental stage [16]. However, *R. collo-cygni* can also colonise the host asymptotically (Fig. 1d; [13]). Expression of *Ramularia* leaf spot symptoms has been linked to host genetics [17–19] but, as with many other Dothideomycete diseases, expression of symptoms has a strong environmental component, in particular responding to light intensity [8, 20–22]. This lifestyle, asymptomatic vertical transmission through

seeds combined with the ability to switch lifestyle when faced with an adverse host environment, is reminiscent of that followed by many plant endophytes [23, 24]. Indeed, it was proposed that *R. collo-cygni* was ancestrally an endophyte and that pathogenicity is a more recent evolutionary phenomenon [25]. Whether changes in pathogen or host genetics or agronomic practices have resulted in the emergence of Ramularia leaf spot as an important pathogen of barley since the late 1990s is currently unclear.

Many Dothideomycetes produce secondary metabolites that are involved in fungal pathogenesis and/or virulence [26]. *R. collo-cygni* produces a number of anthraquinone-derived secondary metabolites called rubellins [27, 28]. Rubellin D, the most widely studied of these compounds, is a light activated, nonhost specific phytotoxin [26]. Rubellins are likely to be synthesised through a polyketide synthase pathway [29], similar to toxin production in other Dothideomycetes [26, 30, 31]. Rubellin D may act by increasing reactive oxygen species (ROS) production resulting in  $\alpha$ -linolenic acid peroxidation, leaf chlorosis and necrosis [27, 28]. How *R. collo-cygni* produces these toxic secondary metabolites and what roles rubellins have in fungal colonisation and disease expression remains to be determined.

To address these and other questions relating to the biology of this fungus and its phylogenetic position within the Dothideomycetes, we have sequenced and assembled the genome of *R. collo-cygni* isolate DK05 Rcc001 (isolated from a susceptible host). We compare the genome to those of other plant pathogens including *Zyoseptoria tritici* [syn. *Mycosphaerella graminicola*], *Pseudocercospora fijiensis* [syn. *Mycosphaerella fijiensis*] and *Dothistroma septosporum* [syn. *Mycosphaerella pini*] from the Mycosphaerellaceae. The high-quality genome of *R. collo-cygni* provides a foundation for future studies aimed at understanding why Ramularia leaf spot has emerged as an important threat to barley production.

## Results and discussion

### Genome assembly and annotation of gene models

The 30.3 Mb *R. collo-cygni* genome was sequenced using a combination of Illumina and Roche 454 FLX technologies to 90-fold coverage and assembled into 576 contigs ranging from 200 bp to 1,386,477 bp in size (Table 1). There were 355 contigs greater than 1 kb in length which had an average GC content of 51.5 % and accounted for 30.1 Mb of the assembly (N50 = 201,222 bp). We predicted 11,617 protein-coding gene models, of which 8514 had transcript evidence from RNAseq analysis of in vitro fungal cultures (see Methods). The average coding sequence length was 1423 bp, with a maximum length of 21,156 bp. Both the

**Table 1** General features of the *Ramularia collo-cygni* isolate DK05 Rcc001 genome assembly

Genome size (bp)	30,300,614
Coverage	90x
Average GC (%)	51.4
Total number of contigs	576
Number of contigs >1 kb	355
N50 (bp) contigs >1 kb	210,222
Max contig size (bp)	1,386,477
Min contig size (bp)	200
Total number of coding sequences (CDS)	11,617
Average length of CDS (bp)	1423
Average coding density <sup>a</sup>	0.546077

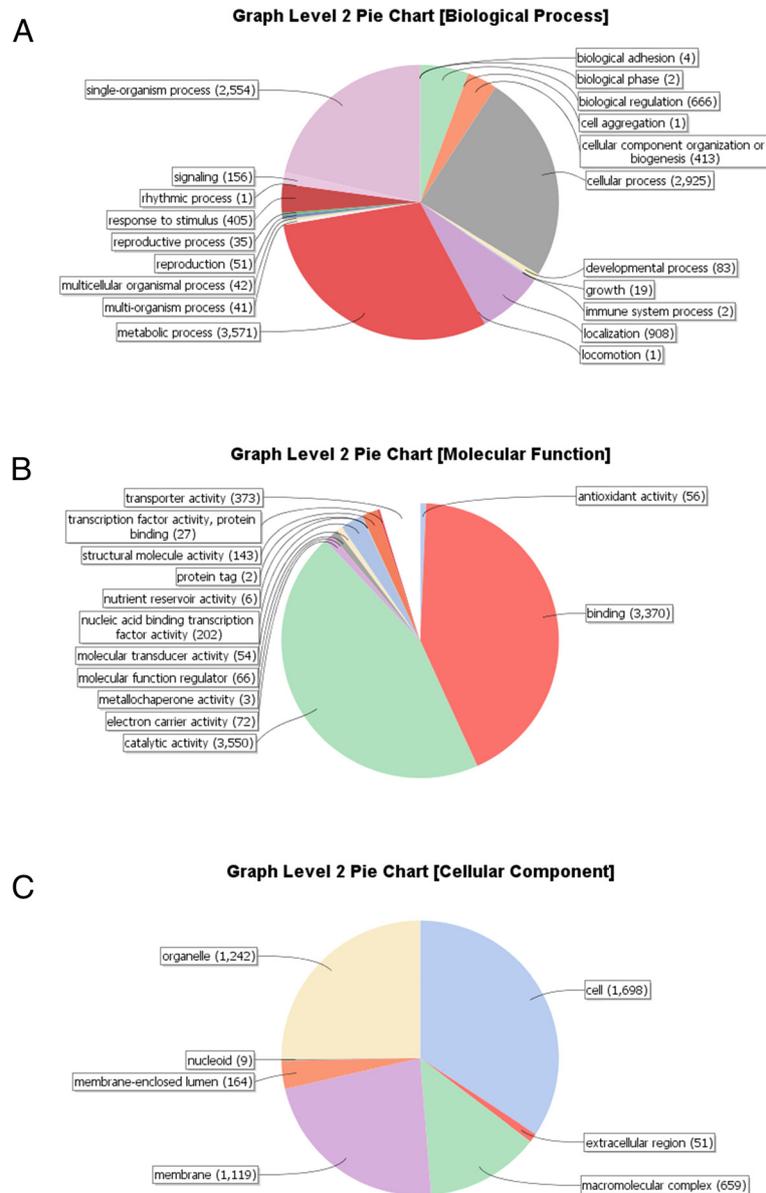
<sup>a</sup>Average coding density = Total number of CDS bases/Total genome bases

estimated *R. collo-cygni* genome size and number of predicted gene models are similar to those of other Dothideomycetes [32] including *Z. tritici* [33] and *D. septosporum* [30]. The genome appears relatively complete, with 94 % of the 248 core eukaryotic gene models in the CEGMA toolkit judged to be complete in the assembly. A genome browser presenting the *R. collo-cygni* genome data can be found at <http://ramularia.org/jbrowse> and the sequence data has been submitted to the European Nucleotide Archive [<http://www.ebi.ac.uk/ena/data/view/PRJEB11432>].

Each predicted gene model was annotated using Blast2GO (Additional file 1) [34, 35]. Most of the top BLASTp matches for the predicted protein set were to *Z. tritici* (36.5 %), *D. septosporum* (15.8 %) and *P. fijiensis* (11.2 %), as expected from their phylogenetic relatedness (Additional file 2). One sixth (1989 gene models; 17.1 %) had no significant matches in the NCBI nr database. For 7442 *R. collo-cygni* gene models we gathered 24,526 level 2 gene ontology (GO) terms which were classified into the categories biological process, molecular function and cellular component (Fig. 2).

### Phylogenetic relationships of *R. collo-cygni* and Dothideomycete fungi

We identified 1026 orthologous protein sets that were largely single-copy across 30 fungal taxa, focusing on Dothideomycetes. Phylogenetic analysis of these alignments clearly positioned *R. collo-cygni* within the order Capnodiales in the Mycosphaerellaceae (Fig. 3). Within the Mycosphaerellaceae *R. collo-cygni* was sister to *Z. tritici*. The other three Mycosphaerellaceae assessed (*P. fijiensis*, *C. fulvum* and *D. septosporum*) formed a sister clade to *R. collo-cygni* plus *Z. tritici*, congruent with previous, less-complete studies which did not include *R. collo-cygni* [30, 32].

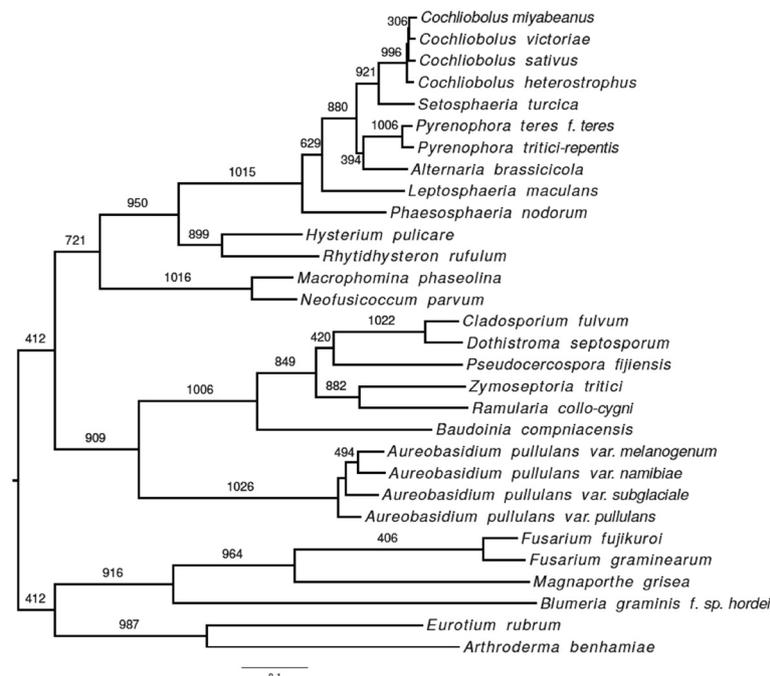


**Fig. 2** Distribution of Blast2GO gene ontology (GO) categories from the putative *Ramularia collo-cygni* gene model annotation. **a** Biological processes, **b** Molecular function, **c** Cellular component

### The *R. collo-cygni* secretome and candidate effector proteins

We identified 1053 genes encoding putative secreted proteins, approximately 9 % of the *R. collo-cygni* proteome, ranging from 45 to 2512 amino acids in length. The predicted *R. collo-cygni* secretome is similar in size to secretomes of other Dothideomycetes such as *Z. tritici* [33, 36] and *D. septosporum* [30]. Most of the predicted secreted proteins (854; 81 %) had significant sequence similarity (BLASTp E-values less than  $1e^{-6}$ ) to proteins from other organisms, with matches to *Z. tritici* (324; 30.7 %), *P. fijiensis* (127; 12 %), *D. septosporum*

(122; 11.6 %) and *Sphaerulina musiva* (94; 8.9 %) most common. Most (59.6 %) of the predicted secreted proteins were annotated with GO terms, the commonest being “oxidation and reduction”, “carbohydrate metabolism” and “proteolysis” in the biological process hierarchy, and “hydrolase activity including hydrolysis of carbohydrates” and “proteins with redox related functions including oxidoreductase and peroxidase activity” in the molecular function hierarchy (Additional file 3). In the *Z. tritici* secretome the protein family domain PF01238, corresponding to peroxidase\_2 family or chloroperoxidase, was abundant [36]. Transcriptome profiling



**Fig. 3** Phylogenetic relationships between *Ramularia collo-cygni* and 29 other fungi with sequenced genomes. Maximum likelihood phylogenetic tree based on a supermatrix analysis of 1026 proteins from 30 taxa (24 Dothideomycetes and six other ascomycetes: see Methods). Branch support is shown as the number of gene trees out of 1026 that supported the presented topology. Bootstrap values are 99–100 for each branch and therefore not shown. Branch length corresponds to a mean estimate of substitutions per site as indicated by the scale bar

during *Z. tritici* infection highlighted up-regulation of chloroperoxidase genes during asymptomatic intercellular growth of the fungus [37]. This family was also found to be expanded in genomes of other plant pathogens within the Mycosphaerellaceae compared to other plant pathogenic fungi [36]. Matches to this domain were found in 21 different proteins in the *R. collo-cygni* secretome (Additional file 3). Most other Mycosphaerellaceae fungi have 15 or more chloroperoxidase genes whereas most other plant pathogenic fungi including Dothideomycetes of the order Pleosporales typically have less than 10 of these genes [36]. This expanded family of chloroperoxidases may play an important role during the endophytic stage of *R. collo-cygni* colonisation.

Many experimentally validated plant pathogen effectors are secreted, cysteine-rich, low molecular weight proteins, termed small secreted proteins (SSP; [38]). A total of 150 proteins from the predicted *R. collo-cygni* secretome matched these criteria (Additional file 4). Effectors are frequently pathogen-specific and just over half (78) of the *R. collo-cygni* SSPs had no significant similarity to proteins found in public databases, suggesting they are unique to this species. Previously reported estimates of the proportion of species-specific SSPs across Dothideomycetes range from 20 to 30 % [32, 39]. The rest of the *R. collo-cygni* SSPs (48 %) had significant similarities to proteins from other organisms, in particular

*Z. tritici* (24; 33.3 %), *S. musiva* (12; 16.7 %) and *P. fijiensis* (10; 13.9 %). Only 18 (12 %) SSPs had predicted PFAM domains, reflecting the typically unknown function of effector proteins. Sixty-seven (45 %) of the predicted SSPs had transcript support from in vitro grown fungal mycelial RNASeq. *R. collo-cygni* SSP gene models that had no transcript support are of particular interest as they may only be expressed *in planta*. Overall, while there appears to be a degree of sequence conservation between secreted proteins of *R. collo-cygni* and other *Mycosphaerella* fungi the function of these putative effector proteins in the biology of the different diseases is currently unknown. Variation between repertoires of putative effector proteins of each species are likely to contribute to the distinct biology of these pathogens. It should also be noted that the analysis of the putative effector content of the *R. collo-cygni* genome may not have been exhaustive. Evidence from the genome of the obligate biotroph *B. graminis* f. sp. *tritici* has indicated the presence of putative effector proteins that do not contain a signal peptide suggesting that they are either non-secreted or secreted through an alternative pathway [40]. Furthermore, plant pathogen effectors are often associated with genomic regions rich in repetitive DNA [41]. Due to the small nature of effector proteins it is possible that the gene annotation process may not have detected them. As such further analysis of the putative effector complement and subsequent characterisation of

species-specific SSPs that are specifically expressed *in planta* may provide further clues to the mechanisms of *R. collo-cygni* pathogenicity.

### Pathogenicity and virulence genes

Fungi use diverse infection strategies for host colonisation and fungal development. We used the Pathogen-Host Interaction database, PHI-base [42, 43], to determine the presence in *R. collo-cygni* of orthologues of pathogenicity genes experimentally confirmed in other fungal species. A total of 1291 *R. collo-cygni* gene models matched 547 PHI-base accessions (Additional file 5). These genes encoded transporters, transcription factors, secondary metabolite biosynthesis enzymes and previously characterised effectors from other pathogens, including three BEC-type effectors from *B. graminis* f. sp. *hordei* [44]. Of particular interest was the identification of 73 *R. collo-cygni* gene models that matched known Dothideomycetes pathogenicity determinants that are associated with toxin production. We identified putative orthologues of proteins involved in biosynthesis of the host-specific HC-toxin (PHI:97; PHI:157) and T-toxin (PHI:2834; PHI:2835; PHI:2836; PHI:2837; PHI:2838; PHI:2839) from *Cochliobolus* spp. and the AK-toxin (PHI:133; PHI:134; PHI:2831; PHI:2832), ACR-toxin (PHI:2608) and ACT-toxin (PHI:2431) produced by *Alternaria alternata* [26] were identified in the *R. collo-cygni* genome. Putative orthologues of genes involved in the biosynthesis (PHI:1046; PHI:1047; PHI:1048; PHI:1049; PHI:1050; PHI:1051) and transport (PHI:141) of the cercosporin toxin produced by *Cercospora* spp. [45] were also identified. Cercosporin is a perylenequinone compound that is a light activated nonhost specific pathogenesis-related toxin [46]. The rubellin toxins produced by *R. collo-cygni* also cause light-activated *in planta* necrosis, fatty acid peroxidation and ROS production [28]. Although to date the only toxins characterised in *R. collo-cygni* are the rubellins, it is probable that this fungus produces a complex arsenal of toxins to facilitate host colonisation and niche exploitation. Determining whether *R. collo-cygni* produces toxins related to cercosporin or toxins characterised in other Dothideomycetes is an important research goal.

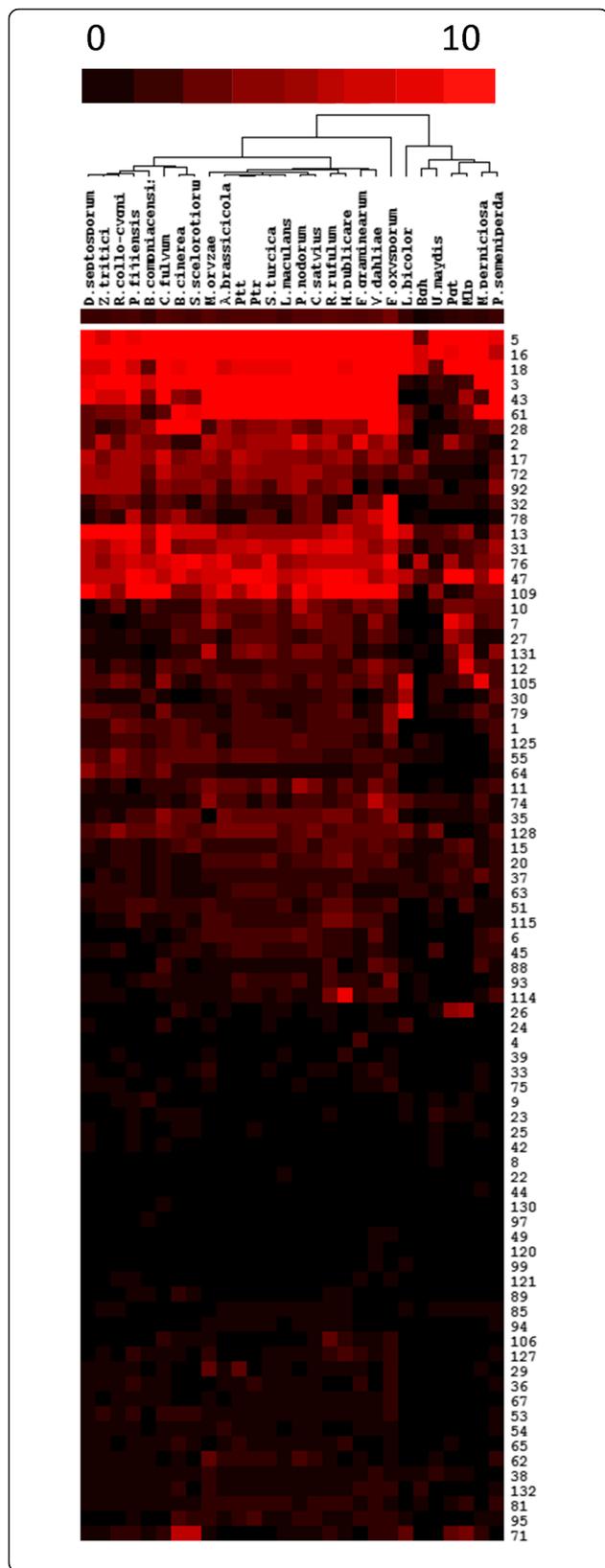
### Carbohydrate-active enzymes

Fungal pathogens need to derive carbohydrates from their hosts and therefore express an array of enzymes capable of metabolising different carbohydrate substrates. The diversity of carbohydrate-active enzymes (CAZymes) can provide insights into the biology of fungal interactions with their specific hosts [47]. *R. collo-cygni* had a total of 520 CAZymes, 226 of which were predicted to be secreted. We identified 223 glycoside hydrolases (GH), 5 polysaccharide lyases (PL), 101

carbohydrate esterases (CE), 55 auxiliary activities (AA), 107 glycosyltransferases (GT) and 29 carbohydrate-binding modules (CBM; Additional file 6). The overall CAZyme complement of *R. collo-cygni* is consistent with that observed in other Dothideomycetes [32, 47]. GH enzymes hydrolyze bonds linking carbohydrates to other molecules [48]. Specific GH family members act on different polysaccharide components of plant cell walls such that the genomic complement of these enzymes can be associated with the trophic habit of fungal pathogens [33, 47]. The GH family complement was compared to that of 27 other fungal species with differing trophic strategies including other Dothideomycetes, Ascomycetes and Basidiomycetes [32]. Cluster analysis of the GH family positioned the Capnodiales in a cluster distinct from the Pleosporales (Fig. 4) in agreement with previous reports [32, 47]. Within the Capnodiales cluster *R. collo-cygni* was placed in a sub-cluster together with the *Mycosphaerella* species (Fig. 4). Further inspection of the GH complement of *R. collo-cygni* highlighted that *Mycosphaerella* species and *R. collo-cygni* have a reduced complement of cellulose-degrading GH enzymes compared to other Dothideomycetes (Figs. 4 and 5). In particular, *R. collo-cygni* has a severely reduced complement of AA9 (formerly known as GH61) copper-dependent, lytic polysaccharide monooxygenase enzymes and no GH6 (endoglucanase and cellobiohydrolase) or GH7 ( $\beta$ -1,4-glucanase, endo- $\beta$ -1,3-1,4-glucanase, reducing end-acting cellobiohydrolase, and chitinase) enzymes. The GH7 family members were present in all of the other fungi examined except the biotrophs *B. graminis* f. sp. *hordei* and *Ustilago maydis*, and the symbiont *L. bicolor* (Fig. 4; Additional file 6).

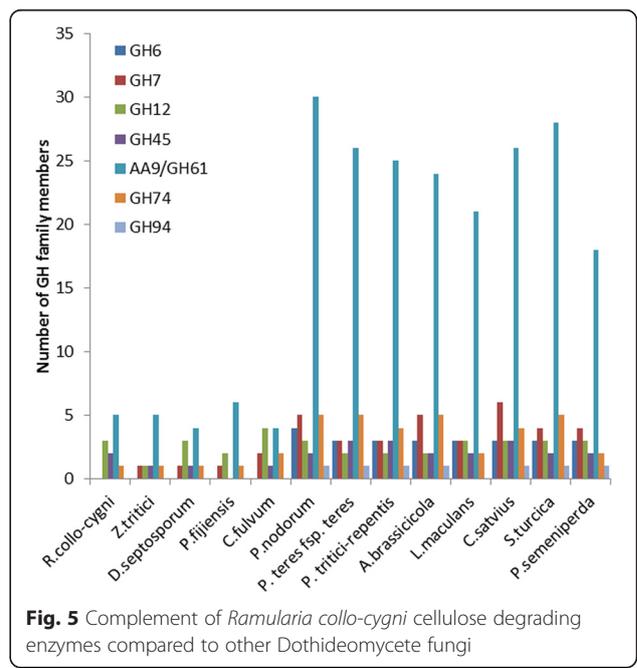
Many economically important fungal pathogens of the genus *Mycosphaerella* have long latent periods of asymptomatic growth before disease occurs [20, 49–51]. As such these fungi need to avoid detection by the host's defence system to allow successful colonisation. *Z. tritici* and *C. fulvum* secrete carbohydrate-binding module 50 (CBM50) domain proteins containing LysM motifs that act to sequester chitin and interfere with chitin-triggered host immunity [52, 53], preventing recognition of the pathogen by the host chitin receptors CERK1 and CEBiP1 [54]. *R. collo-cygni* has at least three CBM50 containing proteins (Table S5) which is in the range for most Dothideomycetes except for *Colletotrichum* spp. where gene expansion of the CBM50 family has been reported [55, 56].

Plant immune systems can recognise and degrade fungal chitin, and some species display active remodelling of the cell wall during invasion. CBM14 and CBM18 family proteins bind chitin [57], and the CBM14-domain effector protein AVR4 from the biotroph fungus *C. fulvum* can bind chitin at the fungal cell wall to prevent its



**Fig. 4** Comparison of *Ramularia collo-cygni* glycoside hydrolase (GH) family complement with 27 other fungal species. GH families and fungal species are hierarchically organised with the heat map indicating the number of members in each family (Black = 0, bright red = >10). Ptt = *Pyrenophora teres* f. sp. *teres*; Ptr = *Pyrenophora tritici-repentis*; Bgh = *Blumeria graminis* f. sp. *hordei*; Pgt = *Puccinia graminis* f. sp. *tritici*; Mlp = *Melampsora larici-populina*

enzymatic degradation by chitinases [58]. No CBM14 family proteins were detected in the *R. collo-cygni* genome. A lack of AVR4-like proteins was also reported for *Z. tritici* where instead the CBM50 containing Mg1LysM and Mg3LysM proteins are able to bind chitin and protect fungi from hydrolytic chitinases [52]. *R. collo-cygni* has six CBM18 family proteins, similar to other *Mycosphaerella* species, but compared to other Dothideomycetes, *R. collo-cygni* and the Capnodiales typically have a contracted CBM18 domain family. Some biotrophic rust fungi and *Colletotrichum graminicola* have been shown to convert chitin in the fungal cell walls of hyphae growing on the leaf surface to chitosan, through the action of chitin deacetylase (CE4) and chitin synthase (GT2), once invasive hyphae penetrate the leaf surface [59], and this may prevent host chitinases from digesting invading hyphae. *R. collo-cygni* has twenty-one GT2 proteins which is at the higher end of the range observed in Dothideomycetes but similar to the GT2 complement of *P. fijiensis* and the extremophilic saprotroph *Baudoinia compniacensis* [60]. *R. collo-cygni* also has six CE4 proteins, the same number found in *Z. tritici*, *D. septosporum*, *P. fijiensis* and *C. fulvum*, but fewer than in any of the Pleosporales fungi (Additional file 6).



**Fig. 5** Complement of *Ramularia collo-cygni* cellulose degrading enzymes compared to other Dothideomycete fungi

The reduced number of plant cell wall degrading enzymes in *R. collo-cygni* coupled with the presence of genes with known roles in avoiding chitin recognition by host defence systems suggests that this species may also utilise “stealth pathogenesis”, as proposed for *Z. tritici* [33]. *R. collo-cygni* is frequently detected in the field in the absence of visible disease [6]. Fungal biomass can build up during asymptomatic development suggesting the fungus is able to avoid recognition by the host [19]. Expression of Ramularia leaf spot symptoms in hosts has a host genetic component [17, 18, 61] but is also modulated by changes in host abiotic stress [8, 12] and development [16]. These features of the pathogenesis of Ramularia leaf spot, combined with the vertical transmission of *R. collo-cygni*, suggest that this species is actually an endophyte that only becomes pathogenic under specific conditions [1, 25]. The closely-related *Z. tritici* is proposed to have evolved from an endophytic ancestor [33], and *R. collo-cygni* may be an endophyte in which pathogenicity has evolved more recently [25]. Determining the roles that CAZymes play during the symptomless and necrotrophic phases will provide valuable insights into the fungal and host cues that trigger *R. collo-cygni* switching from endophytic to pathogenic development.

### Peptidases

Peptidases have multiple functions in plant pathogens including degradation of host defence proteins, signalling and nutrition. *R. collo-cygni* was predicted to have 365 putative peptidases of which 46 % were serine peptidases. Metallopeptidases (87; 23.8 %) and cysteine peptidases (64; 17.5 %) were also highly represented, whereas lower numbers of aspartic (18; 4.9 %) and threonine (20; 5.5 %) peptidases and single glutamic and N9 asparagine peptidases were identified. Five gene models were classified as peptidase inhibitors (Table 2). The distribution of peptidase classes observed in *R. collo-cygni* was similar to that in *Z. tritici*, *P. fijiensis* and *D. septosporum* (Table 2). Ninety-four (25.8 %) peptidases were predicted to be secreted, and most of these were either serine (55; 32.5 %) or metallo-peptidases (21; 24.1 %). Typically serine, metallo- and cysteine peptidases are the most prevalent types in Dothideomycetes whereas asparagine peptidases are relatively uncommon [32, 39].

### Secondary metabolites and rubellin toxin biosynthesis cluster

Many plant pathogenic fungi produce a wide range of secondary metabolites, some of which have important roles in virulence and disease. These are derived from four core biosynthetic origins: polyketide synthases (PKS); non-ribosomal peptide synthases (NRPS); terpene cyclases (TC; syn. terpene synthase) and dimethylallyl tryptophan synthases (DMATS) [26]. *R. collo-cygni* produces nonhost-specific photodynamic anthraquinone toxins, called rubellins [27–29] that are most likely derived from polyketides [29]. Within the *R. collo-cygni* genome 19 PKS, fourteen NRPS and four TC were located (Table 3). No DMATS were identified. *R. collo-cygni* has similar numbers of NRPS and TC genes as other Dothideomycetes. However, the number of PKS in *R. collo-cygni* is nearly double that described for other members of the Capnodiales (Table 3; [30, 32]).

The rubellin toxins produced by *R. collo-cygni* are predicted to be synthesised through a polyketide-derived pathway [29]. Polyketide-derived toxins from other Dothideomycetes, such as the *D. septosporum*-produced dothistromin, are synthesised through pathways similar to that used to produce aflatoxin in *Aspergillus* spp. [22, 31, 62]. Biosynthesis of aflatoxin requires at least 25 enzymes as well as some regulatory proteins, and the genes encoding these are clustered together in a 70 kb region of the *Aspergillus* genome (Fig. 6; [63, 64]). An ancestral core cluster, required to form the initial polyketide product, has been proposed that comprises the genes *AfPksA*, *Affas-1*, *Affas-2* and *Afnor-1*, possibly along with regulatory genes *AfAflR* and *AfAflJ* [65]. Dothistromin has some structural similarity to the aflatoxin intermediate versicolorin B [31] and functional orthologues of aflatoxin biosynthesis genes have been described in *D. septosporum*, although the dothistromin metabolic cluster is fragmented into *DsPksA*, *DsVbsA*, *DsAflR/DsAflJ*, *DsEST*, and *DsVer1* mini-clusters, across a single chromosome (Fig. 6; [22, 30, 31, 66]). Homologs of most dothistromin biosynthesis genes have been found in other Dothideomycete genomes with the complete set identified in *C. fulvum*, a sister species to *D. septosporum*, even though no dothistromin production by this fungus has been reported [30, 67].

**Table 2** Peptidase complement of *Ramularia collo-cygni*, *Zymoseptoria tritici*, *Dothistroma septosporum* and *Pseudocercospora fijiensis*

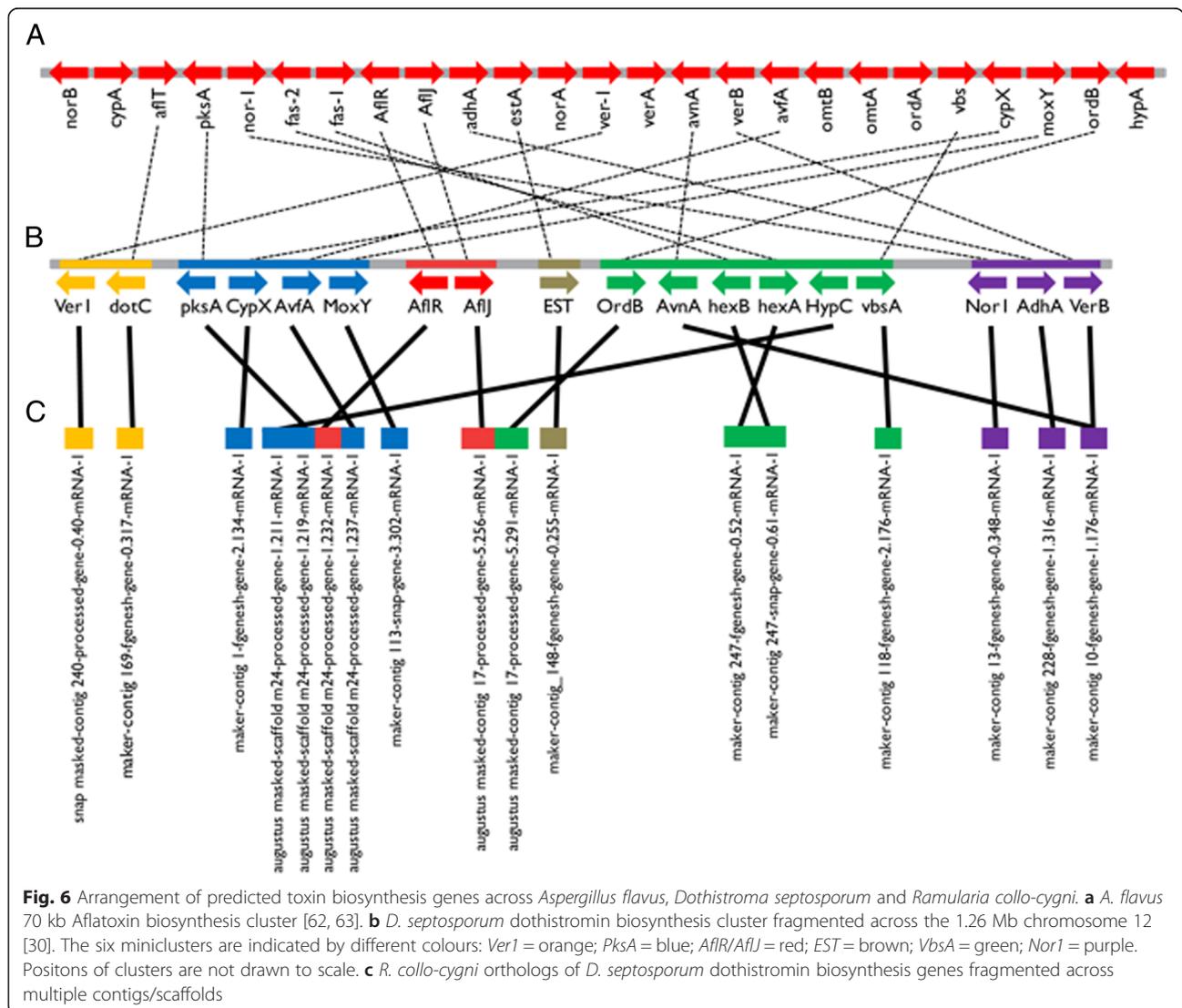
	Aspartic	Cysteine	Glutamic	Inhibitor	Metallo	Asparagine	Serine	Threonine	Total
<i>R. collo-cygni</i>	18	64	1	5	87	1	169	20	365
<i>Z. tritici</i>	22	53	4	7	79	0	155	18	338
<i>D. septosporum</i>	15	49	1	6	75	0	160	19	325
<i>P. fijiensis</i>	15	65	2	6	84	0	173	21	366

Data based on [32]

**Table 3** Comparison of lifestyle and key secondary metabolism genes between *Ramularia collo-cygni* and other selected Dothideomycetes

	Lifestyle	Polyketide synthase (PKS)	Non-ribosomal peptide synthase (NRPS)	Terpene cyclase/synthase (TC/TS)
<i>Ramularia collo-cygni</i>	Endophyte/necrotroph	19	14	4
<i>Zymoseptoria tritici</i> <sup>a</sup>	Hemibiotroph/necrotroph	11	9	5
<i>Dothiostroma septosporum</i> <sup>a</sup>	Hemibiotroph	6	7	7
<i>Pseudocercospora fijiensis</i> <sup>a</sup>	Hemibiotroph	8	11	6
<i>Cladosporium fulvum</i> <sup>a</sup>	Biotroph	10	12	5
<i>Stagonospora nodorum</i> <sup>a</sup>	Necrotroph	19	10	7
<i>Pyrenophora teres</i> f.sp <i>teres</i> <sup>a</sup>	Necrotroph	22	44	11
<i>Pyrenophora repentis-tritici</i> <sup>a</sup>	Necrotroph	17	16	7

<sup>a</sup>Data based on [30] and [32]



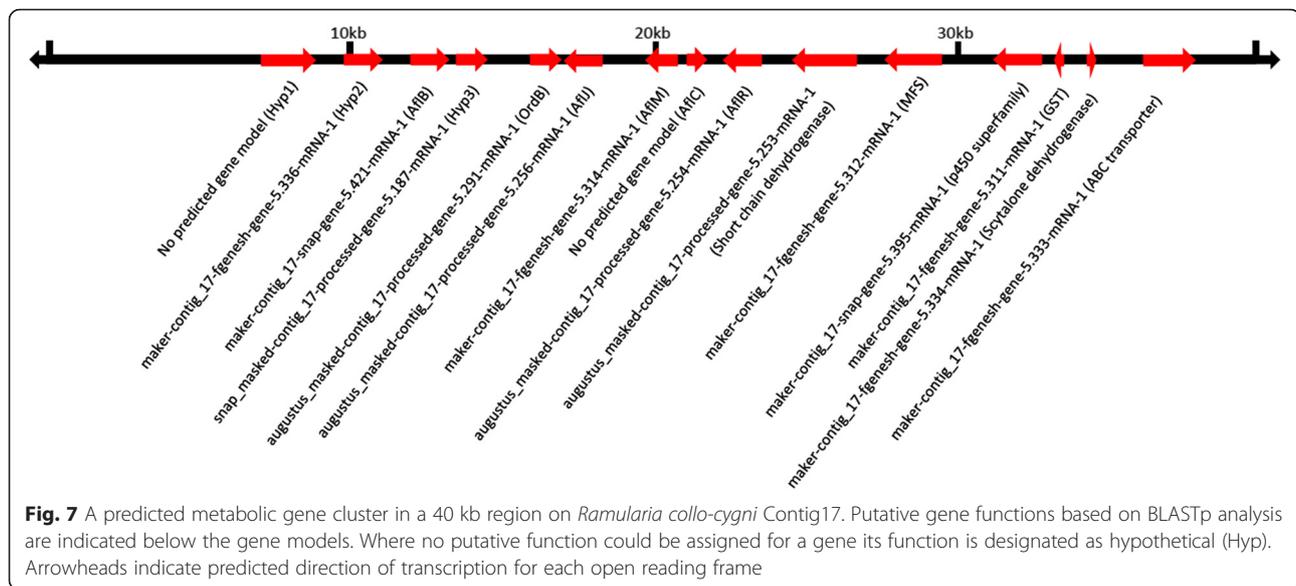
As *R. collo-cygni* produces the polyketide-derived rubellin toxins and due to the close phylogenetic relationship between *R. collo-cygni* and *D. septosporum* (Fig. 3) we used the protein sequences of known dothistromin and aflatoxin biosynthesis genes [30] to identify candidate toxin biosynthesis genes in *R. collo-cygni* (Additional file 7). Homologues of all 18 genes found in the six dothistromin biosynthetic mini-clusters [68] were identified (Additional file 7). However, reciprocal BLASTp analyses indicated that only nine of these proteins were likely to be orthologous to the canonical toxin production loci. The others were identified as orthologues of related enzymes, as has been reported from most other Dothideomycetes [32]. *D. septosporum* *DsAvnA* and *DsVerB* both had the closest similarity to the same *R. collo-cygni* gene model. The dothistromin biosynthetic clusters found in *D. septosporum* were not conserved in *R. collo-cygni* (Fig. 6). Two putative homologs of genes from each of the *DsPksA* and *DsVbsA* clusters appear to co-locate in the *R. collo-cygni* genome. *RccPksA* (a polyketide synthase) and *RccAvfA* (a NAD(P) reductase) were identified on the 229,962 bp scaffold m24 (Fig. 6), whereas *RccHexA* and *RccHexB*, fatty acid synthases from the *DsVbsA* cluster, were located on the 42,586 bp contig247 (Fig. 6). While the *R. collo-cygni* genome has not been assembled to chromosome level, the lack of observed synteny for these loci is telling and suggests that any toxins produced by *R. collo-cygni* are likely to be synthesized through pathways distinct from those in *D. septosporum*.

While *C. fulvum* contains the complete dothistromin biosynthetic pathway, the toxin is not known to be produced by this fungus, and essentially no expression of dothistromin-biosynthesis related transcripts has been observed *in planta* or *in vitro* [30]. We probed expression of the predicted *R. collo-cygni* homologs of dothistromin biosynthesis genes using RT-PCR analysis of *in vitro* and *in planta*. Expression was assessed in RNA extracted from 5, 10, 15 and 20 day old fungal cultures grown in Alkyl Ester broth (AEB) and from barley leaf samples collected from naturally infected field grown plants at two growth stages (GS78 and GS83). Ramularia leaf spot symptoms were clearly visible on plants at both growth stages with no significant difference ( $P = 0.440$ ) in disease levels (Additional file 8) even though there was significantly less green leaf area retention at GS83 ( $P < 0.001$ ; Additional file 8). All of the *R. collo-cygni* homologs of dothistromin biosynthesis genes, including the nine true orthologues and the closest *R. collo-cygni* gene models to remaining *D. septosporum* genes, were expressed in at least one of the *in vitro* time points (Additional file 7). *RccCypX*, *RccAvfA*, *RccMoxY*, *RccAflJ*, *RccOrdB*, *RccHypC* and *RccAdhA* transcripts were expressed *in planta* at GS78 but not at GS83 when

green leaf area retention had declined, whereas *RccPksA*, *RccHexA* and *RccHexB* were expressed at both *in planta* growth stages (Additional file 9). *RccVer1*, *RccdodC*, *RccAflR*, *RccVbsA*, *RccNor1*, *RccAvnA/VerB* and *RccEST1* showed no *in planta* expression at the time points studied. It will be of interest to ascertain whether or not any of these genes are involved in the production of rubellin or other toxins or secondary metabolites that have important biological functions in *R. collo-cygni* development or the expression of disease symptoms.

*AfAflR* regulates transcription of most of the genes in the aflatoxin biosynthetic cluster [62] and the *D. septosporum* *DsAflR* orthologue regulates expression of genes involved in dothistromin biosynthesis [69] suggesting that AflR-like proteins may act as transcriptional regulators of toxin genes across fungal species. *AfAflR* is located in the middle of the aflatoxin biosynthesis gene cluster adjacent to the divergently transcribed *AfAflJ* which has also been shown to be involved in transcriptional regulation of aflatoxin biosynthesis [62]. In *Cercospora nicotianae*, the *AflR*-like *CTB8* and *AflJ*-like *CTB3* are components of the core cercosporin biosynthetic cluster [45]. *R. collo-cygni* homologs of the regulatory gene *DsAflJ* from the dothistromin *AlfR/AflJ* cluster and *DsOrdB*, an oxidoreductase, from the *DsVbsA* cluster were adjacent to each other on Contig17, the largest contig in the assembly (644 kb) containing a putative dothistromin biosynthesis gene homolog (Fig. 7). The 40 kb region surrounding *RccAflJ* and *RccOrdB* on Contig17 contained 13 gene models and two additional open reading frames (ORF), eight of which had significant similarity (BLASTp E-value less than  $1e^{-6}$ ) to proteins from Dothideomycetes (Additional file 10). Four had matches to proteins with predicted functions including a putative ABC transporter, two short chain dehydrogenase/reductases and a scytalone dehydratase, all of which have been reported to have roles in the production of toxins, virulence and/or pathogenicity in Dothideomycetes and other fungi (Additional file 5; [70, 71]).

The remaining loci were most similar to hypothetical proteins from sequenced genomes, and had protein domain matches to short chain dehydrogenase/reductase, DUFF1772, major facilitator superfamily, glutathione S-transferase and p450 superfamily domains. Gene model augustus\_masked-contig17-processed-gene-5.254-mRNA-1 was predicted to encode an *AflR*-like Zn2Cys6 transcription factor containing GAL4-like and AflR domains [62]. The biosynthetic roles of the genes on Contig17 remain to be determined, but all were expressed in barley leaves at GS78 showing Ramularia leaf spot symptoms whereas only the aldoketoreductase-like *AflB1* and MFS-superfamily transporter genes were expressed in diseased tissues at GS83 (Additional file 9). Most of the Contig17 cluster genes were also expressed between 5–20 days during in



vitro fungal growth except the p450, *AflC* and scytalone dehydratase-like genes (Additional file 9). Expression of many of the genes in this cluster in vitro and during disease symptom development prior to excessive leaf senescence may indicate a role for this cluster in fungal development before *R. collo-cygni* enters extensive necrotrophic growth. Functional analysis of the genes within the cluster on Contig17 will provide further insights into their biological role.

The high number of PKS genes in the genome suggests that *R. collo-cygni* may be able to produce other toxins in addition to the characterised rubellins [27–29]. An analysis of gene expression of various *R. collo-cygni* secondary metabolite biosynthesis genes under differing in vitro and *in planta* growth conditions will provide valuable insights into the roles these genes play in the biology of this organism. Experiments are underway to test whether the predicted orthologues of dothistromin/ aflatoxin biosynthesis genes are involved in the production of secondary metabolites including rubellin in *R. collo-cygni*.

## Conclusions

*Ramularia* leaf spot recently emerged as an important disease of barley in temperate regions across the world. The 30.3 Mb genome of *R. collo-cygni* was predicted to contain 11,617 gene models, metrics consistent with the genome size and gene content of other Dothideomycetes. Phylogenetic analysis as well as observed similarities between *R. collo-cygni* and *Z. tritici*, *P. fijiensis* and *D. septosporum* at the gene level support classification of *R. collo-cygni* within the Mycosphaerellaceae [72]. Differences between the genomes of these fungi may hold the key to the unique pathology of *R. collo-cygni*. In

particular identifying the fungal genes involved in the transition from endophytic growth to necrotrophy, the biosynthesis of the rubellin toxins and a further understanding of the genetic structure of *R. collo-cygni* populations [73] are important goals. Projects are currently underway within the *R. collo-cygni* research community to sequence additional genomes from geographically distinct isolates and those collected from non-barley hosts to improve our understanding of how this endophyte has recently established itself as the cause of a newly important foliar disease of barley.

## Methods

### Biological material

*R. collo-cygni* isolate DK05 Rcc001 was isolated in Denmark from the spring barley cv. Braemar, which is highly susceptible to *Ramularia* leaf spot [17], in 2005. Fungal cultures were maintained on potato dextrose agar (PDA, Sigma, Dorset, UK) plates at 16 °C as described previously [12]. Liquid cultures were prepared from an agar plug excised from a PDA plate with seven days fungal growth and placed in to 150 mL AEB [74] supplemented with 5 µg mL<sup>-1</sup> streptomycin. Cultures were incubated at 16 °C under constant agitation at 120 rpm in the dark for 10–12 days. *R. collo-cygni* hyphae were filtered, freeze dried and ground to a fine powder prior to DNA extraction using the Illustra Nucleon PhytoPure Genomic DNA Extraction Kit (GE Healthcare Life Sciences, Little Chalfont, UK), according to the manufacturer's instructions. Genomic DNA was resuspended in 1 % TE buffer (10 mM Tris, 1 mM EDTA- ethylenediaminetetraacetic acid) and subjected to phenol-chloroform-isoamyl alcohol purification. Contaminating RNA was removed by treating the DNA with RiboShredder™ RNase

blend (Cambio Ltd., Cambridge, UK) for 5 h at 37 °C followed by a further round of phenol-chloroform-isoamyl alcohol purification and resuspension of the DNA in 1 % TE. Total RNA was extracted from freeze dried DK05 Rcc001 hyphae scraped from 10 to 12 day old PDA culture plates grown as above using the Trizol (Life Technologies, Paisley, UK) method following the manufacturer's instructions. Contaminating genomic DNA was removed using DNase-free kit as per the manual (Ambion, Leicestershire, UK).

### Genome assembly

The genome and in vitro transcriptome of *R. collo-cygni* were sequenced using a combination of Illumina Genome Analyser IIx and Roche 454 FLX Titanium technologies by the Edinburgh Genomics facility in the University of Edinburgh (<https://genomics.ed.ac.uk/>). Illumina 150 base sequencing generated 2.7 billion bases of genomic data from two libraries of 250 and 350 bp insert sizes and 1.3 billion bases of transcriptome data, and 454 sequencing 31.1 million genome and 3.4 million transcriptome bases (read lengths 100–500 bases). RNASeq data were assembled using the Trinity pathway (<http://trinityrnaseq.github.io/>) using default parameters. The COPE paired end connection tool [75] was used to merge the short insert Illumina reads and a draft assembly generated using the connected short insert library Illumina reads, the raw data from the longer insert library and the 454 data with CLC assembly cell (v.4.0.6 beta). This assembly was then scaffolded using the assembled transcriptome data using SCUBAT (B. Elsworth, G. Koutsopoulos, unpublished). The assembled genome was assessed with REAPR ([76], where all fragment coverage distribution errors that scored above 0.5 were inspected by eye in Tablet [77]). The taxon annotated GC coverage plot approach (<https://github.com/DRL/blobtools-light>) was used to screen the assembly for contaminants (Additional file 11). Contigs with coverage below 9, one tenth of the average coverage, were removed from the assembled genome. Completeness of the draft genome assembly was assessed using CEGMA v.2.4 [78]. The MAKER (<http://www.yandell-lab.org/software/maker.html>) annotation pipeline was used to predict gene models. The final gene models are consensus sequences derived from MAKER using the gene prediction tools snap, genemark, augustus and fgenesh.

### Functional annotation

Protein sequences were compared using BLASTp (E-value threshold less than  $1e^{-6}$ ) against the NCBI nr database and gene ontology (GO) terms assigned using the default settings in Blast2GO (v. 2.8.0; [34, 35]).

### Phylogenetic analysis

Proteomes of 25 Dothideomycetes including 18 plant pathogens and seven saprotrophs were downloaded from the Joint Genome Initiative (<http://genome.jgi-psf.org/>) for phylogenetic analysis. Six representative ascomycete proteomes (two Eurotiomycetes, one Leotiomycete and three Sordariomycetes) were also downloaded as out-group species (Additional file 12; [30, 32, 33, 79–92]). OrthoMCL 2.0 [93] was used to identify gene clusters and 1737 single copy genes for the 31 downloaded fungal genomes plus *R. collo-cygni* were identified using the OrthoMCL output. Single copy genes were identified and aligned using Mafft linsi [94]. The alignments were masked using Zorro [95] with a cutoff value of 4.0 and checked for recombination using PhiPack [96] with default settings. After the recombination test 1251 alignments remained. Two taxa, *Cochliobolus carbonum* and *Cochliobolus heterostrophus* (C4), were removed prior to phylogenetic analyses due to high similarity with *Cochliobolus heterostrophus* (C5) and *Cochliobolus victoriae*, respectively. Two hundred and twenty five genes had several taxa with identical sequences and these were not analysed. Phylogenetic trees were inferred for each gene separately in RAxML [97, 98] using 20 starting trees and model estimation from alignment with PROT-GAMMAUTO. The resulting 1026 gene trees were summarised in a majority rule consensus tree calculated using Consense [99]. Rate of gene evolution was estimated from the branch lengths of the individual gene trees, and the genes were divided into an upper quartile of rapidly evolving genes, a lower quartile of slowly evolving genes, and one partition of medium rate genes. The three partitions were concatenated and analysed using RAxML with 100 rapid bootstraps and ML search from 20 starting trees with model PROTGAMMAUTO for each partition. The individual gene trees were ultrametricised and analysed in DensiTree.

### Prediction of secretome and analysis of small secreted proteins (SSPs) and putative pathogenicity genes

The *R. collo-cygni* secretome was defined as the set of proteins with signal peptides but no transmembrane domain. Signal peptide and transmembrane domain predictions were derived from the InterProScan results of the Blast2GO analysis. SSPs were identified from the *R. collo-cygni* secretome as proteins smaller than 200 amino acids with >2 % cysteine residues. *R. collo-cygni* homologues of experimentally validated pathogenicity genes were identified using PHI-base v. 3.6 (<http://www.phi-base.org/>; [42, 43]). The database was searched using the *R. collo-cygni* gene models using BLASTp with an E-value threshold value of  $1e^{-10}$ . Only PHI-base entries characterised as being associated with reduced virulence, hypervirulence, loss of pathogenicity, a mixed

outcome or predicted effectors were included in the analysis.

#### Carbohydrate active enzymes (CAZymes) annotation

*R. collo-cygni* genes encoding CAZymes were identified using the dbCAN database (<http://csbl.bmb.uga.edu/dbCAN/index.php>) with default settings. The GH content of *R. collo-cygni* was compared with the GH family complement of 27 other fungal species (Table S9; [30, 32, 33, 79, 80, 83–85, 89, 90, 92, 100–107]) using hierarchical clustering. Fungal GH families were clustered using Cluster 3 [108] with a Euclidean distance matrix and complete-linkage clustering. Data were visualized with Treeview v.1.0.13 ([http://www.eisenlab.org/eisen/?page\\_id=42](http://www.eisenlab.org/eisen/?page_id=42)).

#### Peptidases

*R. collo-cygni* peptidases were predicted via BLASTp queries of the MEROPS database (<http://merops.sanger.ac.uk/index.shtml>) using default settings. For comparative purposes the peptidase complements of *Z. tritici*, *D. septosporum* and *P. fijiensis* were also predicted. Proteins were considered peptidases using a threshold E-value of  $1e^{-5}$  [109]. Where multiple peptidases could be assigned to a gene model, the hit with the most significant E-value was selected. *R. collo-cygni* secreted peptidases were identified by cross referencing gene models with the InterProScan scan results from Blast2GO.

#### Secondary metabolite annotation and toxin biosynthesis cluster analysis

PKS, NRPS, TC and DMATS genes were identified using BLASTp searches of protein models against the NCBI nr protein database (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Genes putatively involved in the biosynthesis of toxins were identified using BLASTx and BLASTp searches of the *R. collo-cygni* genome sequence and protein models, respectively, using candidate sequences known to be involved in the biosynthesis of aflatoxin [62, 63] and dothistromin [30]. Contig17 was analysed as a potential gene cluster for additional genes involved in toxin production using the StarORF application (HTML version; Massachusetts Institute of Technology, USA). Putative open reading frames (ORF) were identified on Contig17 and the protein sequences screened against the NCBI nr protein database. Each putative ORF was confirmed by BLASTp analysis back against the *R. collo-cygni* gene models.

Gene specific primers were designed for candidate toxin biosynthesis genes using Primer3 (<http://primer3.ut.ee/>). Transcript expression was assessed during *in vitro* and *in planta* *R. collo-cygni* growth. *In vitro* expression was measured in fungal hyphae grown in AEB cultures harvested at 5, 10, 15, 20 days and *in planta* expression was measured in *Ramularia* leaf spot infected spring barley flag -1

(F-1) leaves collected from a field trial experiment (see Field experiment sampling). Total RNA was extracted as described above and cDNA was synthesised from 1  $\mu$ g of total RNA using the Superscript III system (Invitrogen, Carlsbad, CA, USA) and target expression assessed using end point RT-PCR. cDNA was diluted 20-fold in dH<sub>2</sub>O and 5  $\mu$ L used to amplify each target using gene specific primers (Additional file 13) and the Hotstar Master mix. Target sequences were amplified using a TProfessional Standard Thermocycler (Biometra, Göttingen, Germany) under the following cycling conditions: 15 min enzyme hot start activation at 95 °C followed by 35 cycles of 1 min denaturation at 94 °C, 1 min primer annealing at 60 °C and 1 min extension at 72 °C and a final extension cycle of 10 min at 72 °C.

#### Field experiment sampling

Leaves of the spring barley cv. Concerto exhibiting *Ramularia* leaf spot symptoms were collected from a 10 m  $\times$  2 m plot of a field trial sown at the Bush Estate, Midlothian, Scotland in 2014. Herbicide and fertiliser treatments in the trial followed local practice with the sampled plot treated with Prothioconazole (0.44 L ha<sup>-1</sup>, Proline 275) at GS21-35 and GS41. F-1 leaf samples were collected at growth stages GS78 and GS83 and *Ramularia* leaf spot infection and green leaf area retention of each sample recorded as a percentage of leaf area covered. Leaf samples were stored at -80 °C for fungal gene expression analysis as described above (see Secondary metabolite annotation and toxin biosynthesis cluster analysis section).

#### Additional files

**Additional file 1: Table S1.** Blast2GO annotation of *Ramularia collo-cygni* gene models. (XLSX 1285 kb)

**Additional file 2: Figure S1.** *Ramularia collo-cygni* gene model top species hits from BLASTp analysis. (TIF 1644 kb)

**Additional file 3: Table S2.** Predicted secretome of *Ramularia collo-cygni*. (XLSX 152 kb)

**Additional file 4: Table S3.** Small secreted protein complement of *Ramularia collo-cygni* genome. (XLSX 43 kb)

**Additional file 5: Table S4.** Putative *Ramularia collo-cygni* pathogenicity and virulence genes identified by analysis of PHI-base. (XLSX 96 kb)

**Additional file 6: Table S5.** Carbohydrate active enzyme (CAZyme) complement of *Ramularia collo-cygni* genome compared to selected other fungi. (XLSX 43 kb)

**Additional file 7: Table S6.** Homologs of dothistromin and aflatoxin biosynthesis genes in *Ramularia collo-cygni* genome. (XLSX 12 kb)

**Additional file 8: Figure S2.** Data from naturally infected *Ramularia* leaf spot spring barley field trials. A. *Ramularia* leaf spot levels on cv. Concerto at GS78 and GS83 (% of the total leaf area). B. Green leaf area retention of *Ramularia* leaf spot infected samples at GS78 and BS83. (TIF 902 kb)

**Additional file 9: Table S7.** RT-PCR analysis of *Ramularia collo-cygni* homologs of dothistromin biosynthesis genes and predicted secondary metabolism genes on Contig17. (XLSX 13 kb)

**Additional file 10: Table S8.** BLASTp top hits for putative metabolic cluster genes on Contig17. (XLSX 11 kb)

**Additional file 11: Figure S3.** Taxon annotated GC coverage plot of *Ramularia collo-cygni* genome assembly. (TIF 1542 kb)

**Additional file 12: Table S9.** Details of fungal genomes used for comparative genomics in this study. (XLSX 17 kb)

**Additional file 13: Table S10.** RT-PCR primer details. (XLSX 13 kb)

## Abbreviations

AA, auxiliary activities; AEB, alkyl ester broth; CAZymes, carbohydrate-active enzymes; CBM, carbohydrate-binding modules; CE, carbohydrate esterases; DMATS, dimethylallyl tryptophan synthases; DUF, domain of unknown function; GH, glycoside hydrolase; GO, gene ontology; GS, growth stage; GT, glycosyl-transferases; NRPS, non-ribosomal peptide synthases; ORF, open reading frame; PDA, potato dextrose agar; PKS, polyketide synthases; PL, polysaccharide lyases; ROS, reactive oxygen species; SSP, small secreted protein; TC, terpene cyclases

## Acknowledgements

This research was supported by grant-in-aid from the Scottish Government RESAS (GMcG, JMF, FD, MK, AM, JMT, KG, LJP FB and NDH) and by a Danish government grant 'BRCC- Restricting the pathogenic effect on barley of the endophytic fungus *Ramularia collo-cygni*' (ES and AM). AA, KK, JJR and KHK are supported by strategic funding from the United Kingdom Biotechnology and Biological Sciences Research Council (BBSRC) including the Institute Strategic Programme Grant 20:20 Wheat® (BB/J/00426X/1). KHK and the PHI-base project are supported by the BBSRC PhytoPath grant (BB/K020056/1) and the National Capability grant (BB/J/004383/1). Edinburgh Genomics is partly supported through core grants from NERC (R8/H10/56) and BBSRC (BB/J004243/1). We thank Steve Goodwin and Craig Coleman for access to the *P. fijiensis* and *P. semeniperda* sequence data and the Joint Genome Institute for access to sequence data produced by the US Department of Energy Joint Genome Institute <http://www.jgi.doe.gov/> in collaboration with the user community.

## Funding

This research was funded by a Scottish Government RESAS grant, a Danish government Strategiske Forskningsråd grant, Biotechnology and Biological Sciences Research Council grants BB/K020056/1, BB/J/004383/1, BB/J/00426X/1, BB/J004243/1 and Natural Environment Research Council grant R8/H10/56. None of these funding bodies were involved in the design of the study nor the collection, analysis, and interpretation of data and in writing the manuscript.

## Availability of data and materials

The data sets supporting the results in this article are available at the European Nucleotide Archive [<http://www.ebi.ac.uk/ena/data/view/PRJEB11432>] and a genome browser presenting the *R. collo-cygni* genome data can be found at <http://ramularia.org/jbrowse>.

## Authors' contributions

GMcG, JMF, FD, MK, AM, JMT, KG, LJP, NDH, FB, MB, UT, ES, AA, JJR, KK, KHK designed and performed the experiments. AA, UT, ES, GMcG performed the genome assembly and annotation. GMcG wrote the manuscript with input from all authors. All authors have read and approved the manuscript prior to submission.

## Competing interests

The authors declare that they have no competing interests.

## Consent for publication

Not applicable

## Ethics approval and consent to participate

Not applicable

## Author details

<sup>1</sup>Crop Protection Team, Crop and Soil Systems Group, SRUC, West Mains Road, Edinburgh EH9 3JG, UK. <sup>2</sup>Department of Computational and Systems Biology, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK.

<sup>3</sup>Department of Plant Biology and Crop Science, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK. <sup>4</sup>Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3TF, UK. <sup>5</sup>Edinburgh Genomics, The University of Edinburgh, Edinburgh EH9 3JT, UK. <sup>6</sup>Present address: Forest Research, Alice Holt Lodge, Farnham, Surrey GU10 4LH, UK. <sup>7</sup>Present address: Syngenta, Jealott's Hill International Research Centre, Bracknell, Berkshire RG42 6EY, UK.

Received: 9 November 2015 Accepted: 12 July 2016

Published online: 09 August 2016

## References

- Havis ND, Brown JKM, Clemente G, Frei P, Jedryczka M, Kaczmarek J, Kaczmarek M, Matusinsky P, McGrann GRD, Pereyra S, Piotrowska M, Sghyer H, Tellier A, Hess M. *Ramularia collo-cygni* - an emerging pathogen of barley crops. *Phytopathology*. 2015;105:895–904.
- Cavara F. Über einige parasitische Pilze auf dem Getreide. *Zeitschrift Pflanzenkrankheit*. 1893;3:16–26.
- Sutton BC, Waller JM. Taxonomy of *Ophiocladium hordei*, causing leaf lesions on Triticale and other Graminaeae. *Trans Br Mycol Soc*. 1988;90:55–61.
- Crous PW, Aptroot A, Kang JC, Braun U, Wingfield MJ. The genus *Mycosphaerella* and its anamorphs. *Stud Mycol*. 2000;45:107–21.
- Crous PW, Kang JC, Braun U. A phylogenetic redefinition of anamorph genera in *Mycosphaerella* based on ITS rDNA sequence and morphology. *Mycologia*. 2001;93:1081–101.
- Havis ND, Oxley SJP, Piper SR, Langrell SRH. Rapid nested PCR-based detection of *Ramularia collo-cygni* direct from barley. *FEMS Microbiol Lett*. 2006;256:217–23.
- Frei P, Gindro K, Richter H, Schuerch S. Direct-PCR detection and epidemiology of *Ramularia collo-cygni* associated with barley necrotic leaf spots. *J Phytopathol*. 2007;155:281–8.
- Makepeace JC, Havis ND, Burke JI, Oxley SJP, Brown JKM. A method of inoculating barley seedlings with *Ramularia collo-cygni*. *Plant Pathol*. 2008; 57:991–9.
- Taylor JMG, Paterson LJ, Havis ND. A quantitative real-time PCR assay for the detection of *Ramularia collo-cygni* from barley (*Hordeum vulgare*). *Lett Appl Microbiol*. 2010;50:493–9.
- Thirugnanasambandam A, Wright KM, Havis N, Whisson SC, Newton AC. Agrobacterium-mediated transformation of the barley pathogen *Ramularia collo-cygni* with fluorescent marker tags and live tissue imaging of infection development. *Plant Pathol*. 2011;60:929–37.
- Huss H. The biology of *Ramularia collo-cygni*. Meeting the challenges of barley blights. *Proceedings of the second international workshop on barley leaf blights 2004, Aleppo, Syria: ICARDA*; 2002. p 321–328.
- Peraldi A, Griffe LL, Burt C, McGrann GRD, Nicholson P. *Brachypodium distachyon* exhibits compatible interactions with *Oculimacula* spp. and *Ramularia collo-cygni*, providing the first pathosystem model to study eyespot and ramularia leaf spot diseases. *Plant Pathol*. 2014;63:554–62.
- Havis ND, Nyman M, Oxley SJP. Evidence for seed transmission and symptomless growth of *Ramularia collo-cygni* in barley (*Hordeum vulgare*). *Plant Pathol*. 2014;63:929–36.
- Stabentheiner E, Minihofer T, Huss H. Infection of barley by *Ramularia collo-cygni*: scanning electron microscopic investigations. *Mycopathologia*. 2009; 168:135–43.
- Kaczmarek M, Piotrowska MJ, Fontaine JM, Gorniak K, McGrann GRD, Armstrong A, Wright DM, Newton AC, Havis ND. Infection strategy of *Ramularia collo-cygni* and development of ramularia leaf spot on barley and alternative graminaceous hosts. *Plant Pathol*. 2016. doi: 10.1111/ppa.12552.
- Schützendübel A, Stadler M, Wallner D, von Tiedemann A. A hypothesis on physiological alterations during plant ontogenesis governing susceptibility of winter barley to ramularia leaf spot. *Plant Pathol*. 2008;57:518–26.
- McGrann GRD, Stavrinides A, Russell J, Corbitt MM, Booth A, Chartrain L, Thomas WT, Brown JKM. A trade off between *mlo* resistance to powdery mildew and increased susceptibility of barley to a newly important disease, Ramularia leaf spot. *J Exp Bot*. 2014;65:1025–37.
- McGrann GRD, Steed A, Burt C, Goddard R, Lachaux C, Bansal A, Corbitt M, Gorniak K, Nicholson P, Brown JKM. Contribution of the drought tolerance-related *Stress-responsive NAC1* transcription factor to resistance of barley to Ramularia leaf spot. *Mol Plant Pathol*. 2015;16:201–9.

19. McGrann GRD, Steed A, Burt C, Nicholson P, Brown JKM. Differential effects of lesion mimic mutants in barley on disease development by facultative pathogens. *J Exp Bot*. 2015;66:3417–28.
20. Keon J, Antoniw J, Carzaniga R, Deller S, Ward JL, Baker JM, Beale MH, Hammond-Kosack K, Rudd JJ. Transcriptional adaptation of *Mycosphaerella graminicola* to programmed cell death (PCD) of its susceptible wheat host. *Mol Plant Microbe Interact*. 2007;20:178–93.
21. Churchill AC. *Mycosphaerella fijiensis*, the black leaf streak pathogen of banana: progress towards understanding pathogen biology and detection, disease development, and the challenges of control. *Mol Plant Pathol*. 2011;12:307–28.
22. Bradshaw RE. Dothistroma (red-band) needle blight of pines and the dothistromin toxin: a review. *For Pathol*. 2004;34:163–85.
23. Schulz B, Boyle C. The endophytic continuum. *Mycol Res*. 2005;109:661–86.
24. Rodriguez R, Redman R. More than 400 million years of evolution and some plants still can't make it on their own: plant stress tolerance via fungal symbiosis. *J Exp Bot*. 2008;59:1109–14.
25. Salamati, S, and Reitan, L. *Ramularia collo-cygni* on spring barley, an overview of its biology and epidemiology. Proceedings 1st European Ramularia Workshop 2006. Göttingen, Germany: ICARDA; 2006. p 19–35.
26. Muria-Gonzalez MJ, Chooi Y-H, Breen S, Solomon PS. The past, present and future of secondary metabolite research in the Dothideomycetes. *Mol Plant Pathol*. 2015;16:92–107.
27. Heiser I, Sachs E, Liebermann B. Photodynamic oxygen activation by rubellin D, a phytotoxin produced by *Ramularia collo-cygni* (Sutton et Waller). *Physiol Mol Plant Pathol*. 2003;62:29–36.
28. Heiser I, Hess M, Schmidtku KU, Vogler U, Miethbauer S, Liebermann B. Fatty acid peroxidation by rubellin B, C and D, phytotoxins produced by *Ramularia collo-cygni* (Sutton et Waller). *Physiol Mol Plant Pathol*. 2004;64:135–43.
29. Miethbauer S, Haase S, Schmidtku KU, Guenther W, Heiser I, Liebermann B. Biosynthesis of photodynamically active rubellins and structure elucidation of new anthraquinone derivatives produced by *Ramularia collo-cygni*. *Phytochemistry*. 2006;67:1206–13.
30. de Wit PJ, van der Burgt A, Okmen B, Stergiopoulos I, Abd-Elsalam KA, Aerts AL, Bahkali AH, Beenen HG, Chettri P, Cox MP. The genomes of the fungal plant pathogens *Cladosporium fulvum* and *Dothistroma septosporium* reveal adaptation to different hosts and lifestyles but also signatures of common ancestry. *PLoS Genet*. 2012;8:e1003088.
31. Bradshaw RE, Zhang S. Biosynthesis of dothistromin. *Mycopathologia*. 2006;162:201–13.
32. Ohm RA, Feau N, Henrissat B, Schoch CL, Horwitz BA, Barry KW, Condon BJ, Copeland AC, Dhillon B, Glaser F, Hesse CN, Kosti I, LaButti K, Lindquist EA, Lucas S, Salamov AA, Bradshaw RE, Ciuffetti L, Hamelin RC, Kema GH, Lawrence C, Scott JA, Spatafora JW, Turgeon BG, de Wit PJ, Zhong S, Goodwin SB, Grigoriev IV. Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen Dothideomycetes fungi. *PLoS Pathog*. 2012;8:e1003037.
33. Goodwin SB, Ben M'Barek S, Dhillon B, Wittenberg AHJ, Crane CF, Hane JK, Foster AJ, Van der Lee TAJ, Grimwood J, Aerts A, Antoniw J, Bailey A, Bluhm B, Bowler J, Bristow J, van der Burgt A, Canto-Canche B, Churchill ACL, Conde-Ferreaez L, Cools HJ, Coutinho PM, Csukai M, Dehal P, De Wit P, Donzelli B, van de Geest HC, van Ham RCHJ, Hammond-Kosack KE, Henrissat B, Kilian A, Kobayashi AK, Koopmann E, Kourmpetis Y, Kuzniar A, Lindquist E, Lombard V, Maliepaard C, Martins N, Mehrabi R, Nap JPH, Ponomarenko A, Rudd JJ, Salamov A, Schmutz J, Schouten HJ, Shapiro H, Stergiopoulos I, Torriani SFF, Tu H, de Vries RP, Waalwijk C, Ware SB, Wiebenga A, Zwieters LH, Oliver RP, Grigoriev IV, Kema GHJ. Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. *PLoS Genet*. 2011;7:e1002070.
34. Conesa A, Gotz S. Blast2GO: a comprehensive suite for functional analysis in plant genomics. *Int J Plant Genomics*. 2008;2008:619832.
35. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005;21:3674–6.
36. do Amaral AM, Antoniw J, Rudd JJ, Hammond-Kosack KE. Defining the predicted protein secretome of the fungal wheat leaf pathogen *Mycosphaerella graminicola*. *PLoS One*. 2012;e49904.
37. Rudd JJ, Kanyuka K, Hassani-Pak K, Derbyshire M, Andongabo A, Devonshire J, Lysenko A, Saqi M, Desai NM, Powers SJ, Hooper J, Ambrosio L, Bharti A, Farmer A, Hammond-Kosack KE, Dietrich RA, Courbot M. Transcriptome and metabolite profiling of the infection cycle of *Zymoseptoria tritici* on wheat reveals a biphasic interaction with plant immunity involving differential pathogen chromosomal contributions and a variation on the hemibiotrophic lifestyle definition. *Plant Physiol*. 2015;167:1158–85.
38. Kamoun S. A catalogue of the effector secretome of plant pathogenic oomycetes. *Annu Rev Phytopathol*. 2006;44:41–60.
39. Lowe RG, Cassin A, Grandaubert J, Clark BL, Van de Wouw AP, Rouxel T, Howlett BJ. Genomes and transcriptomes of partners in plant-fungal-interactions between canola (*Brassica napus*) and two *Leptosphaeria* species. *PLoS One*. 2014;9:e103098.
40. Wicker T, Oberhaensli S, Parlange F, Buchmann JP, Shatalina M, Roffler S, Ben-David R, Doležel J, Šimková H, Schulze-Lefert P. The wheat powdery mildew genome shows the unique evolution of an obligate biotroph. *Nat Genet*. 2013;45:1092–6.
41. Haas BJ, Kamoun S, Zody MC, Jiang RHY, Handsaker RE, Cano LM, Grabherr M, Kodira CD, Raffaele S, Torto-Alalibo T, Borkzart TO, Ah-Fong AMV, Alvarado L, Anderson VL, Armstrong MR, Avrova A, Baxter L, Beynon J, Boevink PC, Bollmann SR, Bos JIB, Bulone V, Cai GH, Cakir C, Carrington JC, Chawner M, Conti L, Costanzo S, Ewan R, Fahlgren N, Fischbach MA, Fugelstad J, Gilroy EM, Gnerre S, Green PJ, Grenville-Briggs LJ, Griffith J, Grunwald NJ, Horn K, Horner NR, Hu CH, Huitema E, Jeong DH, Jones AME, Jones JDG, Jones RW, Karlsson EK, Kunjeti SG, Lamour K, Liu ZY, Ma LJ, MacLean D, Chibucos MC, McDonald H, McWalters J, Meijer HJG, Morgan W, Morris PF, Munro CA, O'Neill K, Ospina-Giraldo M, Pinzon A, Pritchard L, Ramsahoye B, Ren QH, Restrepo S, Roy S, Sadanandom A, Savidor A, Schornack S, Schwartz DC, Schumann UD, Schwessinger B, Seyer L, Sharpe T, Silvar C, Song J, Studholme DJ, Sykes S, Thines M, van de Vondervoort PJJ, Phuntumart V, Wawra S, Weide R, Win J, Young C, Zhou SG, Fry W, Meyers BC, van West P, Ristaino J, Govers F, Birch PRJ, Whisson SC, Judelson HS, Nussbaum C. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature*. 2009;461:393–8.
42. Urban M, Pant R, Raghunath A, Irvine AG, Pedro H, Hammond-Kosack KE. The Pathogen-Host Interactions database (PHI-base): additions and future developments. *Nucleic Acids Res*. 2015;43(Database issue):D645–55.
43. Urban M, Irvine AG, Cuzick A, Hammond-Kosack KE. Using the Pathogen-Host Interactions database (PHI-base) to investigate plant pathogen genomes and genes implicated in virulence. *Front Plant Sci*. 2015;6:605. doi: 10.3389/fpls.2015.00605.
44. Pliego C, Nowara D, Bonciani G, Gheorghe DM, Xu R, Surana P, Whigham E, Nettleton D, Bogdanove AJ, Wise RP, Schweizer P, Bindschedler LV, Spanu PD. Host-induced gene silencing in barley powdery mildew reveals a class of ribonuclease-like effectors. *Mol Plant Microbe Interact*. 2013;26:633–42.
45. Chen H, Lee M, Daub ME, Chung K. Molecular analysis of the cercosporin biosynthetic gene cluster in *Cercospora nicotianae*. *Mol Microbiol*. 2007;64:755–70.
46. Upchurch RG, Walker DC, Rollins JA, Ehrenshaft M, Daub ME. Mutants of *Cercospora kikuchii* altered in cercosporin synthesis and pathogenicity. *Appl Environ Microbiol*. 1991;57:2940–5.
47. Zhao Z, Liu H, Wang C, Xu JR. Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. *BMC Genomics*. 2013;14:274.
48. Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*. 2014;42:D490–5.
49. Thomma BP, Van Esse HP, Crous PW, de Wit PJ. *Cladosporium fulvum* (syn. *Passalora fulva*), a highly specialized plant pathogen as a model for functional studies on plant pathogenic *Mycosphaerellaceae*. *Mol Plant Pathol*. 2005;6:379–93.
50. Donzelli BGG, Churchill AC. A quantitative assay using mycelial fragments to assess virulence of *Mycosphaerella fijiensis*. *Phytopathology*. 2007;97:916–29.
51. Kabir MS, Ganley RJ, Bradshaw RE. Dothistromin toxin is a virulence factor in dothistroma needle blight of pines. *Plant Pathol*. 2015;64:225–34.
52. Marshall R, Kombrink A, Motteram J, Loza-Reyes E, Lucas J, Hammond-Kosack KE, Thomma BPHJ, Rudd JJ. Analysis of two in planta expressed LysM effector homologs from the fungus *Mycosphaerella graminicola* reveals novel functional properties and varying contributions to virulence on wheat. *Plant Physiol*. 2011;156:756–69.
53. de Jonge R, Van Esse HP, Kombrink A, Shinya T, Desaki Y, Bours R, van der Krol S, Shibuya N, Joosten MH, Thomma BP. Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. *Science*. 2010;329:953–5.
54. Lee WS, Rudd JJ, Hammond-Kosack KE, Kanyuka K. *Mycosphaerella graminicola* LysM effector-mediated stealth pathogenesis subverts

- recognition through both CERK1 and CEBiP homologues in wheat. *Mol Plant Microbe Interact.* 2014;27:236–43.
55. Gan P, Ikeda K, Irieda H, Narusaka M, O'Connell RJ, Narusaka Y, Takano Y, Kubo Y, Shirasu K. Comparative genomic and transcriptomic analyses reveal the hemibiotrophic stage shift of *Colletotrichum* fungi. *New Phytol.* 2013; 197:1236–49.
  56. O'Connell RJ, Thon MR, Hacquard S, Amyotte SG, Kleemann J, Torres MF, Damm U, Buiate EA, Epstein L, Alkan N. Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. *Nat Genet.* 2012;44:1060–5.
  57. Sánchez-Vallet A, Mesters JR, Thomma BP. The battle for chitin recognition in plant-microbe interactions. *FEMS Microbiol Rev.* 2014;39:171–83.
  58. van den Burg HA, Spronk CA, Boeren S, Kennedy MA, Vissers JP, Vuisster GW, de Wit PJ, Vervoort J. Binding of the AVR4 elicitor of *Cladosporium fulvum* to chitotriose units is facilitated by positive allosteric protein-protein interactions. *J Biol Chem.* 2004;279:16786–96.
  59. El Gueddari NE, Rauchhaus U, Moerschbacher BM, Deising HB. Developmentally regulated conversion of surface-exposed chitin to chitosan in cell walls of plant pathogenic fungi. *New Phytol.* 2002;156:103–12.
  60. Scott JA, Untereiner WA, Ewaze JO, Wong B, Doyle D. *Baudoinia*, a new genus to accommodate *Torula compniacensis*. *Mycologia.* 2007;99:592–601.
  61. Saville RJ, Gosman N, Burt CJ, Makepeace J, Steed A, Corbitt M, Chandler E, Brown JKM, Boulton MI, Nicholson P. The 'Green Revolution' dwarfing genes play a role in disease resistance in *Triticum aestivum* and *Hordeum vulgare*. *J Exp Bot.* 2012;63:1271–83.
  62. Yu J. Current understanding on aflatoxin biosynthesis and future perspective in reducing aflatoxin contamination. *Toxins.* 2012;4:1024–57.
  63. Yu J, Chang PK, Ehrlich KC, Cary JW, Bhatnagar D, Cleveland TE, Payne GA, Linz JE, Woloshuk CP, Bennett JW. Clustered pathway genes in aflatoxin biosynthesis. *Appl Environ Microbiol.* 2004;70:1253–62.
  64. Bhatnagar D, Ehrlich KC, Cleveland TE. Molecular genetic analysis and regulation of aflatoxin biosynthesis. *Appl Microbiol Biotechnol.* 2003;61:83–93.
  65. Cary JW, Ehrlich KC. Aflatoxigenicity in *Aspergillus*: molecular genetics, phylogenetic relationships and evolutionary implications. *Mycopathologia.* 2006;162:167–77.
  66. Chettri P, Ehrlich KC, Cary JW, Collemare J, Cox MP, Griffiths SA, Olson MA, de Wit PJ, Bradshaw RE. Dothistromin genes at multiple separate loci are regulated by AfIR. *Fungal Genet Biol.* 2013;51:12–20.
  67. Kabir MS, Ganley RJ, Bradshaw RE. The hemibiotrophic lifestyle of the fungal pine pathogen *Dothistroma septosporum*. *For Pathol.* 2015;45:190–202.
  68. Schwelm A, Bradshaw RE. Genetics of dothistromin biosynthesis of *Dothistroma septosporum*: an update. *Toxins.* 2010;2:2680–98.
  69. Bradshaw RE, Slot JC, Moore GG, Chettri P, Wit PJ, Ehrlich KC, Ganley AR, Olson MA, Rokas A, Carbone I. Fragmentation of an aflatoxin-like gene cluster in a forest pathogen. *New Phytol.* 2013;198:525–35.
  70. Inderbitzin P, Asvarak T, Turgeon BG. Six new genes required for production of T-toxin, a polyketide determinant of high virulence of *Cochliobolus heterostrophus* to maize. *Mol Plant Microbe Interact.* 2010;23:458–72.
  71. Stergiopoulos I, Zwiers LH, De Waard MA. The ABC transporter MgAt4r4 is a virulence factor of *Mycosphaerella graminicola* that affects colonization of substomatal cavities in wheat leaves. *Mol Plant Microbe Interact.* 2003;16:689–98.
  72. Walters DR, Havis ND, Oxley SJP. *Ramularia collo-cygni*: the biology of an emerging pathogen of barley. *FEMS Microbiol Lett.* 2008;279:1–7.
  73. Piotrowska MJ, Ennos RA, Fontaine JM, Burnett FJ, Kaczmarek M, Hoebe PN. Development and use of microsatellite markers to study diversity, reproduction and population genetic structure of the cereal pathogen *Ramularia collo-cygni*. *Fungal Genet Biol.* 2016;87:64–71.
  74. Skinner W, Bailey A, Renwick A, Keon J, Gurr S, Hargreaves J. A single amino acid substitution in the iron-sulphur protein subunit of succinate dehydrogenase determines resistance to carboxin in *Mycosphaerella graminicola*. *Curr Genet.* 1998;34:393–8.
  75. Liu B, Yuan J, Yiu SM, Li Z, Xie Y, Chen Y, Shi Y, Zhang H, Li Y, Lam TW. COPE: an accurate k-mer-based pair-end reads connection tool to facilitate genome assembly. *Bioinformatics.* 2012;28:2870–4.
  76. Hunt M, Kikuchi T, Sanders M, Newbold C, Berriman M, Otto TD. REAPR: a universal tool for genome assembly evaluation. *Genome Biol.* 2013;14:R47.
  77. Milne I, Stephen G, Bayer M, Cock PJ, Pritchard L, Cardle L, Shaw PD, Marshall D. Using Tablet for visual exploration of second-generation sequencing data. *Brief Bioinform.* 2013;14:193–202.
  78. Parra G, Bradnam K, Korf I. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics.* 2007;23:1061–7.
  79. Hane JK, Lowe RG, Solomon PS, Tan KC, Schoch CL, Spatafora JW, Crous PW, Kodira C, Birren BW, Galagan JE. Dothideomycete-plant interactions illuminated by genome sequencing and EST analysis of the wheat pathogen *Stagonospora nodorum*. *Plant Cell.* 2007;19:3347–68.
  80. Ellwood SR, Liu Z, Syme RA, Lai Z, Hane JK, Keiper F, Moffat CS, Oliver RP, Friesen TL. A first genome assembly of the barley fungal pathogen *Pyrenophora teres f. teres*. *Genome Biol.* 2010;11:R109.
  81. Blanco-Ulate B, Rolshausen P, Cantu D. Draft genome sequence of *Neofusicoccum parvum* isolate UCR-NP2, a fungal vascular pathogen associated with grapevine cankers. *Genome Announc.* 2013;1:e00339–13. doi:10.1128/genomeA.00339-13.
  82. Burmester A, Shelest E, Glockner G, Heddergott C, Schindler S, Staib P, Heidel A, Felder M, Petzold A, Szfranski K. Comparative and functional genomics provide insights into the pathogenicity of dermatophytic fungi. *Genome Biol.* 2011;12:R7.
  83. Condon BJ, Leng Y, Wu D, Bushley KE, Ohm RA, Otilar R, Martin J, Schackwitz W, Grimwood J, MohdZainudin N. Comparative genome structure, secondary metabolite, and effector coding capacity across *Cochliobolus* pathogens. *PLoS Genet.* 2013;9:e1003233.
  84. Cuomo CA, Guldener U, Xu JR, Trail F, Turgeon BG, Di Pietro A, Walton JD, Ma LJ, Baker SE, Rep M. The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. *Science.* 2007;317:1400–2.
  85. Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, Thon M, Kulkarni R, Xu JR, Pan H. The genome sequence of the rice blast fungus *Magnaporthe oryzae*. *Nature.* 2005;434:980–6.
  86. Gostincar C, Ohm RA, Kogej T, Sonjak S, Turk M, Zajc J, Zalar P, Grube M, Sun H, Han J. Genome sequencing of four *Aureobasidium pullulans* varieties: biotechnological potential, stress tolerance, and description of new species. *BMC Genomics.* 2014;15:549.
  87. Islam MS, Haque MS, Islam MM, Emdad EM, Halim A, Hossain QM, Hossain MZ, Ahmed B, Rahim S, Rahman MS. Tools to kill: genome of one of the most destructive plant pathogenic fungi *Macrophomina phaseolina*. *BMC Genomics.* 2012;13:493.
  88. Kis-Papo T, Weig AR, Riley R, Persoh D, Salamov A, Sun H, Lipzen A, Wasser SP, Rambold G, Grigoriev IV. Genomic adaptations of the halophilic Dead Sea filamentous fungus *Eurotium rubrum*. *Nat Commun.* 2014;5:3745.
  89. Manning VA, Pandelova I, Dhillon B, Wilhelm LJ, Goodwin SB, Berlin AM, Figueroa M, Freitag M, Hane JK, Henrissat B. Comparative genomics of a plant-pathogenic fungus, *Pyrenophora tritici-repentis*, reveals transduplication and the impact of repeat elements on pathogenicity and population divergence. *G3.* 2013;3:41–63.
  90. Rouxel T, Grandaubert J, Hane JK, Hoede C, van de Wouw AP, Couloux A, Dominguez V, Anthouard V, Bally P, Bourras S, Cozijnsen AJ, Ciuffetti LM, Degraeve A, Dilmaghani A, Duret L, Fudal I, Goodwin SB, Gout L, Glaser N, Linglin J, Kema GH, Lapalu N, Lawrence CB, May K, Meyer M, Ollivier B, Poulain J, Schoch CL, Simon A, Spatafora JW, Stachowiak A, Turgeon BG, Tyler BM, Vincent D, Weissenbach J, Amselem J, Quesneville H, Oliver RP, Wincker P, Balesdent MH, Howlett BJ. Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by Repeat-Induced Point mutations. *Nat Commun.* 2011;2:202.
  91. Wiemann P, Sieber CM, von Bargen KW, Studt L, Niehaus EM, Espino JJ, Huß K, Michiels CB, Albermann S, Wagner D, Bergner SV, Connolly LR, Fischer A, Reuter G, Kleigrew K, Bald T, Wingfield BD, Ophir R, Freeman S, Hippler M, Smith KM, Brown DW, Proctor RH, Münsterkötter M, Freitag M, Humpf HU, Guldener U, Tudzynski B. Deciphering the cryptic genome: genome-wide analyses of the rice pathogen *Fusarium fujikuroi* reveal complex regulation of secondary metabolism and novel metabolites. *PLoS Pathog.* 2013;9:e1003475.
  92. Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, Stuber K, van Themaat EVL, Brown JKM, Butcher SA, Gurr SJ, Lebrun MH, Ridout CJ, Schulze-Lefert P, Talbot NJ, Ahmadinejad N, Ametz C, Barton GR, Benjdia M, Bidzinski P, Bindschedler LV, Both M, Brewer MT, Cadle-Davidson L, Cadle-Davidson MM, Collemare J, Cramer R, Frenkel O, Godfrey D, Harriman J, Hoede C, King BC, Klages S, Kleemann J, Knoll D, Koti PS, Kreplak J, Lopez-Ruiz FJ, Lu XL, Maekawa T, Mahanil S, Micali C, Milgroom MG, Montana G, Noir S, O'Connell RJ, Oberhaensli S, Parlange F, Pedersen C, Quesneville H, Reinhardt R, Rott M, Sacristan S, Schmidt SM, Schon M, Skamnioti P, Sommer H, Stephens A, Takahara H, Thordal-Christensen H, Vigouroux M,

- Wessling R, Wicker T, Panstruga R. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science*. 2010; 330:1543–6.
93. Li L, Stoeckert CJ, Roos DS. OrthoMCL: identification of orthologue groups for eukaryotic genomes. *Genome Res*. 2003;13:2178–89.
  94. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. 2002;30:3059–66.
  95. Wu M, Chatterji S, Eisen JA. Accounting for alignment uncertainty in phylogenomics. *PLoS One*. 2012;7:e30288.
  96. Bruen TC, Philippe H, Bryant D. A simple and robust statistical test for detecting the presence of recombination. *Genetics*. 2006;172:2665–81.
  97. Ott M, Zola J, Aluru S, Stamatakis A, ACM. Large-scale maximum likelihood-based phylogenetic analysis on the IBM BlueGene/L. In: *Acm/leee Sc07 Conference*, vol. Article No. 4. 2007. p. 466–76.
  98. Stamatakis A. RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. 2006;22:2688–90.
  99. Felsenstein J. PHYLIP (Phylogeny Inference Package) version 3.6. Seattle: Department of Genome Sciences, University of Washington; 2005.
  100. Martin F, Aerts A, Ahrén D, Brun A, Danchin EG, Duchaussoy F, Gibon J, Kohler A, Lindquist E, Pereda V, Salamov A, Shapiro HJ, Wuyts J, Blaudez D, Buée M, Brokstein P, Canbäck B, Cohen D, Courty PE, Coutinho PM, Delaruelle C, Detter JC, Deveau A, DiFazio S, Duplessis S, Fraissinet-Tachet L, Lucic E, Frey-Klett P, Fourrey C, Feussner I, Gay G, Grimwood J, Hoegger PJ, Jain P, Kilaru S, Labbé J, Lin YC, Legué V, Le Tacon F, Marmeisse R, Melayah D, Montanini B, Muratet M, Nehls U, Niculita-Hirzel H, Oudot-Le Secq MP, Peter M, Quesneville H, Rajashekar B, Reich M, Rouhier N, Schmutz J, Yin T, Chalot M, Henrissat B, Kües U, Lucas S, Van de Peer Y, Podila GK, Polle A, Pukkila PJ, Richardson PM, Rouzé P, Sanders IR, Stajich JE, Tunlid A, Tuskan G, Grigoriev IV. The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature*. 2008;452:88–92.
  101. Amselem J, Cuomo CA, Van Kan JA, Viaud M, Benito EP, Couloux A, Coutinho PM, de Vries RP, Dyer PS, Fillinger S. Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. *PLoS Genet*. 2011;7:e1002230.
  102. Duplessis S, Cuomo CA, Lin YC, Aerts A, Tisserant E, Veneault-Fourrey C, Joly DL, Hacquard S, Amselem J, Cantarel BL, Chiu R, Coutinho PM, Feau N, Field M, Frey P, Gelhaye E, Goldberg J, Grabherr MG, Kodira CD, Kohler A, Kuees U, Lindquist EA, Lucas SM, Mago R, Mauceli E, Morin E, Murat C, Pangilinan JL, Park R, Pearson M, Quesneville H, Rouhier N, Sakthikumar S, Salamov AA, Schmutz J, Selles B, Shapiro H, Tanguay P, Tuskan GA, Henrissat B, Van de Peer Y, Rouze P, Ellis JG, Dodds PN, Schein JE, Zhong S, Hamelin RC, Grigoriev IV, Szabo LJ, Martin F. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proc Natl Acad Sci U S A*. 2011;108:9166–71.
  103. Kamper J, Kahmann R, Bolker M, Ma LJ, Brefort T, Saville BJ, Banuett F, Kronstad JW, Gold SE, Muller O, Perlin MH, Wösten HA, de Vries R, Ruiz-Herrera J, Reynaga-Peña CG, Snetselaar K, McCann M, Pérez-Martín J, Feldbrügge M, Basse CW, Steinberg G, Ibeas JI, Holloman W, Guzman P, Farman M, Stajich JE, Sentandreu R, González-Prieto JM, Kennell JC, Molina L, Schirawski J, Mendoza-Mendoza A, Grellinger D, Münch K, Rössel N, Scherer M, Vranes M, Ladendorf O, Vincon V, Fuchs U, Sandrock B, Meng S, Ho EC, Cahill MJ, Boyce KJ, Klose J, Klosterman SJ, Deelstra HJ, Ortiz-Castellanos L, Li W, Sanchez-Alonso P, Schreier PH, Häuser-Hahn I, Vaupel M, Koopmann E, Friedrich G, Voss H, Schlüter T, Margolis J, Platt D, Swimmer C, Gnirke A, Chen F, Vysotskaia V, Mannhaupt G, Güldener U, Münsterkötter M, Haase D, Oesterheld M, Mewes HW, Mauceli EW, DeCaprio D, Wade CM, Butler J, Young S, Jaffe DB, Calvo S, Nusbaum C, Galagan J, Birren BW. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature*. 2006;444:97–101.
  104. Klosterman SJ, Subbarao KV, Kang S, Veronese P, Gold SE, Thomma BPHJ, Chen Z, Henrissat B, Lee YH, Park J. Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens. *PLoS Pathog*. 2011;7:e1002137.
  105. Ma LJ, Van Der Does HC, Borkovich KA, Coleman JJ, Daboussi MJ, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature*. 2010;464:367–73.
  106. Mondego JM, Carazzolle MF, Costa GG, Formighieri EF, Parizzi LP, Rincones J, Cotomacci C, Carraro DM, Cunha AF, Carrer H. A genome survey of *Moniliophthora perniciosa* gives new insights into Witches' Broom Disease of cacao. *BMC Genomics*. 2008;9:548.
  107. Soliai MM, Meyer SE, Udall JA, Elzinga DE, Hermansen RA, Bodily PM, Hart AA, Coleman CE. De novo genome assembly of the fungal plant pathogen *Pyrenophora semeniperda*. *PLoS One*. 2014;9:e87045.
  108. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A*. 1998;95:14863–8.
  109. Rawlings ND, Waller M, Barrett AJ, Bateman A. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res*. 2014; 42:D503–9.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

