Genome sequence and spore germination associated transcriptome analysis of Corynespora cassiicola from cucumber

Shigang Gao
Shanghai Academy of Agricultural Sciences https://orcid.org/0000-0001-8165-8821
Rong Zeng
Shanghai Academy of Agricultural Sciences
Yujuan Suo
Shanghai Academy of Agricultural Sciences
Lihui Xu
Shanghai Academy of Agricultural Sciences
Zhiwei Song
Shanghai Academy of Agricultural Sciences
Ping Gao
Shanghai Academy of Agricultural Sciences
Fu Ming Dai (fumingdai@163.com)
https://orcid.org/0000-0001-5139-5303

Research article

Keywords: Corynespora cassiicola, Cucumber, Genome sequence, Virulence-associated gene, Spore germination, RNA-seq

Posted Date: November 14th, 2019

DOI: https://doi.org/10.21203/rs.2.17308/v1

Abstract

Corynespora cassiicola as a necrotrophic plant pathogenic ascomycetes fungus can infect hundreds of species of plants, and also rarely cause human disease. The pathogen infects cucumber and causes cucumber target spot, which has given rise to great yield loss of cucumber in China recently. Genome sequence and spore germination associated transcriptome analysis will contribute to the understanding of the molecular mechanism of pathogenicity and spore germination of C. cassiicola .Results Firstly, we reported the draft genome sequences of a cucumber-sampled C. cassiicola isolate HGCC with high virulence. Although being conspecific, HGCC had distinct difference with a rubber-sampled isolate (CCP) and a human-sampled isolate (UM591) in genome sequences. The proportion of secreted proteins was 7.4% in HGCC. 28.6% of HGCC predicted genes were highly homologous to experimentally proven virulence-associated genes, which was close to that in CCP, UM591 and some plant fungal pathogens, but far more than 21.9% in Phaeosphaeria nodorum and 19.6% in Botrytis cinerea . Thousands of putative virulence-associated genes in various pathways or families were identified in HGCC. Secondly, a global view of the transcriptome of C. cassiicola spores during germination was evaluated using RNA sequencing (RNA-seq). A total of 3,288 differentially expressed genes (DEGs) were identified. The majority of KEGG annotated DEGs were involved in metabolism, genetic information processing, cellular processes, organismal system, human diseases and environmental information processing.Conclusions These results not only facilitated the exploration of the molecular pathogenic mechanism of C. cassiicola to cucumber and the understanding of molecular and cellular processes during spore germination, but also laid the foundation for the disease control.

Background

The cucumber target spot caused by *C. cassiicola* (BerK & Curt) Wei has given rise to tremendous yield loss of cucumber in China recently [1]. More importantly, *C. cassiicola* as a necrotrophic parasitic fungus can infect more than 500 species of plants besides cucumber including tomato, eggplant, tobacco, rubber, cotton, soybean and balsam pear, *et al*. and cause emerging plant spot diseases [2], which have rise to a major disease recently from a minor disease for the last half century. Additionally, *C. cassiicola* is also a human opportunistic pathogen to rarely infect human in suitable condition and cause fungal keratitis [3] and subcutaneous phaeohyphomycosis [4]. Therefore, the pathogenic mechanism of *C. cassiicola* against plants should be well studied to effectively control diseases caused by the pathogen.

For the past few years, the researches for the pathogenic mechanism of *C. cassiicola* mainly focused on biological characteristics, pathogenicity differentiation, cloning of virulence-associated genes, *et al*. It was reported that *C. cassiicola* isolates sampled from cucumbers could grow at from 10 to 35℃ and the optimum temperature was about 30℃ [5]. Spore of cucumber pathogen *C. cassiicola* could be germinated from one end or both ends of the spore at 25-30℃ with above 90% of relative humidity, and the germination rate in water drop was the highest. *C. cassiicola* was able to invade into cucumber leaves mainly directly or via stomata [6]. *C. cassiicola* showed pathogenic and genetic variation among isolates sampled in different host plants, proving that the intraspecific strains of *C. cassiicola* showed high host

specialization [2, 7]. 64 *C. cassiicola* isolates from perilla, cucumber, tomato, aubergine and sweet pepper in Japan were divided into seven pathogenicity groups (PG1-PG7) [8]. Cassiicolin, a small secreted glycoprotein, was an important effector of *C. cassiicola*, and contained six different cassiicolin isoforms such as *Cas1*, *Cas2*, *Cas3*, *Cas4*, *Cas5* and *Cas6* in different *C. cassiicola* isolates sampled from various hosts and geographical origins [9-11]. The aggressive abilities of isolates were related to the types of isoform, and the isolates carrying *Cas1* gene showed the most aggressive on rubbers. Additionally, some isolates with no *Cas* gene also generated moderate symptoms on rubber leaves, showing that uncharacterized effectors existed in the *C. cassiicola* [10].

Like other filamentous fungal pathogens, *C. cassiicola* needs multiple pathogenic factors such as cutinase, cell wall degrading enzymes [12] and cytomembrane and cell inclusion degrading enzymes besides toxin (cassiicolin) to successfully invade into host plants and cause disease, which are transported or regulated by multiple pathways of virulence-associated genes including mitogen-activated protein kinase (MAPK), Ca and cAMP signal pathways *et al* [13]. So far, virulence-associated genes of *C. cassiicola* were rarely cloned and functionally characterized except for two MAPK genes *CCk1* [14] and *CMP1* [15], cassiicolin encoded gene *Cas* [16, 17], which was far from thoroughly understanding the pathogenic mechanism of *C. cassiicola*. Thus, a great number of virulence-associated genes remain to be identified, cloned and functionally characterized.

Fungal conidium being involved in the reproduction is the main type of fungal asexual spores, and also is the main form of inoculation and infection. Resting conidia undergo germination and form sporeling hyphae or thlli under suitable conditions such as humidity and temperature [18]. During this process, the progressive reduction of conidia hydrophobicity results in conidia swelling with isotropic growth and subsequent polarized growth being characterized as germ tube formation [19]. Conidium germination is usually essential for the fungal aggressiveness and colonization, which needs the involving of a series of genes with different biochemical activities [20-22]. Therefore, the spore germination associated gene expression analysis will contribute to the understanding of the molecular mechanism of spore germination and pathogenicity, which is not reported in *C. cassiicola* so far.

Genome sequencing is a high throughput means to identify functional genes being combined with homologous match against functional database or conserved domain search [13, 23]. Furthermore, genome sequences accelerate the clone of genes and their function characterization, so it is imperative that the genome sequence of *C. cassiicola* is globally analyzed. Although the genome sequencing of two *C. cassiicola* isolates CCP from rubber and UM591 from the contact lens of a patient with keratomycosis were completed [24, 25], it was still difficult to characterize the gene function in cucumber-sampled isolates due to high genetic variation among isolates from different hosts. RNA sequencing (RNA-seq) is a effective technique to analyze the expression of a great many of genes and wildly used to identify DEGs between different treatments. Therefore, firstly, we presented the draft genome sequence of a cucumber-sampled *C. cassiicola* isolate (HGCC) with high virulence to cucumber, and comparatively analyzed its genome with another two *C. cassiicola* isolates (CCP from rubber, UM591 from human) and other plant pathogenic ascomycetes fungi in multiple families or pathways of virulence associated genes in this

study. Secondly, we studied the relative transcriptional levels of genes during the spore germination of *C. cassiicola* HGCC using RNA-seq. Research results would provide a great deal of information for revealing the molecular mechanism of pathogenicity and spore germination of *C. cassiicola*.

Results

HGCC isolate features

HGCC isolate with high virulence to cucumber was selected for de novo sequencing of *C. cassiicola* sampled from cucumber. By being sub-cultured on PDA plates at 25℃ in dark, HGCC produced a layer of fluffy aerial mycelium with whitish gray when young and dark green when old (Figure 1A), which was easily peel off. The clubbed conidia were varied in length (from 20 to 120 μm) and shape with one to nine septa, which could germinate at one end or two ends in sterile distilled water (Figure 1B). After being inoculated on sensitive and resistant cucumber cultivars artificially, HGCC caused typical symptom of cucumber target spot disease in the sensitive cultivar Biyu, but it hardly infected the resistant cultivar Shenqing-1 (Figure 1C and D).

Genome sequencing and general features

HGCC genome was de novo sequenced (184 × Coverage) using Illumina Hiseq X-Ten. 54,580,316 high quality reads were assembled into 1,032 scaffolds (N50: 500 kb) with 42.7 Mb of genome size, slightly less than CCP (44.8 Mb) (JGI: 1019537) and greater than UM591 (41.4 Mb) (GenBank: JAQF00000000.1) (Table 1). HGCC genome encoded 15,678 genes close to CCP (15,614) and slightly more than UM591 (15,388) through the prediction of coding sequence (CDS) prediction (Table 1). However, the number of secreted proteins in HGCC (1,166) were less than CCP (1,216) and UM591 (1,182). The proportions of secreted proteins in the three isolates were 7.4% for HGCC, 7.8% for CCP and 7.7% for UM591, which were similar to other phytopathogenic ascomycetes fungi (7-10%) [23].

Interspecific genome-wide phylogeny

HGCC genome had 93.1% and 92.3% amino acid sequence identity with CCP and UM591, and 46.1-58.4% with other phytopathogenic fungi such as *C. lunata* (58.3%), *B. maydis* (58.3%), *S. turcica* (58.3%), *P. nodorum* (58.2%), *P. tritici-repentis* (57.7%), *C. zeae-maydis* (48.4%), *A. flavus* (48%), *B. cinerea* (47.1%), *F. graminearum* (46.4%) and *M. oryzae* (46.1%), respectively. >70% of HGCC genes had >90% amino acid sequence identity with CCP and UM591, far more than 55% between two *C. lunata* isolates CX-3 from maize and m118 from sorghum, respectively [23]. 13,672, 13,701 and 13,715 homologous core genes were screened by reciprocal blast analysis in HGCC, CCP and UM591, of which 1,335, 1,330 and 1,401 were specific to the 10 selected pathogenic fungi (Figure 2A), respectively. Additionally, HGCC, CCP and UM591 had 7.3% (1,138), 6.6% (1,035) and 6.7% (1,037) specific genes compared to each other, of which 6.7% (1,047), 6.3% (961) and 6.3% (971) were specific to the 10 selected pathogenic fungi, respectively (Figure 2A and Table 2). It was suggested that the three *C. cassiicola* isolates had distinct differences in genome, although HGCC had high amino acid sequence identity with CCP and UM591.

12,225 gene encoding proteins were classified into 4,070 conserved protein families in HGCC by Pfam matches with profile hidden Markov models, slightly more than 4,003 families containing 12,238 proteins in CCP and 4,002 families containing 11,918 proteins in UM591. Glycoside hydrolase, pectate lyase, cutinase and cellulose were important pathogenic factors in phytopathogenic fungi that richly existed in *C. cassiicola* (Additional file 1: Table S1). *C. cassiicola* had family expansions in fungal specific transcription factor (P = 0.00451), transporter (P = 0.00005), major facilitator superfamily (P = 0.00002), ATP-binding cassette (ABC) superfamily (P = 0.00985), cytochrome P450s (P = 0.00001), G-protein coupled receptors (P =0.00336), protein kinases (P =0.00008), proteases (P = 0.00004), glycoside hydrolase (P = 0.00035) and pectate lyases (P = 0) compared to *P. nodorum*, *P. tritici-repentis*, *S. turcica*, *B. maydis*, *C. lunata*, *C. zeae-maydis*, *B. cinerea* and *M. oryzae*, *F. graminearum*. These protein families were expected to play important roles in the fungal survival in varied adverse environments.

In order to mine potential virulence-associated genes, Blastp searches of three *C. cassiicola* genomes were conducted against the pathogen-host interaction (PHI) database. 28.6% of predicted genes of HGCC were matched with PHI database at a *E*-value of $1e\text{-}10^5$ and putatively related to the PHI, which was close to 29.0% in CCP, 28.7% in UM591, 25.3% in *P. tritici-repentis*, 28.4% in *S. turcica*, 26.7% in *B. maydis*, 30.8% in *C. lunata*, 25.8% in *C. zeae-maydis*, 32.6% in *A. flavus*, 25.1% in *M. oryzae*, but far more than 21.9% in *P. nodorum* and 19.6% in *B. cinerea*.

An interspecific phylogenomic tree of *C. cassiicola* and the 10 other plant fungi was constructed based on concatenated amino acid sequences of 2,344 core proteins (Figure 2B). The tree showed *C. cassiicola* of Pleosporales order had the genetic affinity with other Pleosporales order fungi, followed by *C. zeae-maydis* of Capnodiales order of Dothideomycetes class, *A. flavus* of Eurotiomycetes class, *B. cinerea* of Leotiomycetes class, and *M. oryzae* and *F. graminearum* of Sordariomycetes class. In addition, it was showed in the tree that *C. cassiicola* speciation had occurred before the speciation of the other Pleosporales order fungi, suggesting that *C. cassiicola* speciation had occurred before the speciation of the Pleosporales order. These results were similar to the result of David Lopez [24].

Pathogenic signal pathway

G-protein-coupled receptors (GPCRs) transduce external environmental signals by way of heterotrimeric G proteins into secondary messengers to regulate gene expression and subsequent cellular response [26]. They are required in plant recognition and pheromone/nutrient sensing of plant pathogenic fungi [27]. Pth11, as one GPCR of *Magnaporthe grisea*, mediates appressorium differentiation and fungal pathogenicity [28], and its homologues exists in other plant pathogenic fungi. HGCC genome contains 181 GPCR-like genes and 77 Pth11-like GPCRs, which is not only more than 165/61 in CCP and 169/55 in UM591, but also far more than 81-156/21-58 in selected 10 other plant pathogenic fungi (Additional file 1: Table S2). 56.4% of GPCRs-like genes and 80.5% of Pth11-like GPCRs are identified as PHI associated genes in HGCC, far more than 28.6% of PHI percentage in HGCC genome. G-protein alpha subunit is a important component of heterotrimeric G protein complex [29], which can activate downstream effectors and function as fungal pathogenicity [30, 31]. In HGCC genome, three G-protein alpha subunit

(HGCC_9368, HGCC_6623 and HGCC_511) were identified, and both of them were both PHI-associated genes. HGCC_9368, HGCC_6623 and HGCC_511 showed high amino acid identities with a G protein beta subunit of *Pseudocercospora fijiensis* (GenBank: XP_007925889.1, 93%) and two G-protein alpha subunits of *Stemphylium lycopersici* (GenBank: KNG46771.1, 95%; GenBank: KNG45504.1, 85%), respectively.

MAPK, cAMP and Ca signaling pathways were main virulence-associated signal pathways, and controlled by a series of protein kinase [13]. Three MAPK pathways of *Saccharomyces cerevisiae*, FUS3/KSS1, Mpk1 and Hog1, were well studied and highly conserved in other fungi [13]. Based on homologous searches of HGCC genome against known MAPK, cAMP or Ca pathways associated genes in *S. cerevisiae*, 48, 12 and 23 genes in HGCC, high homologous with MAPK, cAMP and Ca signal pathways genes of *S. cerevisiae*, were screened, respectively (Additional file 1: Table S3, S4 and S5). 156 protein kinases were identified in HGCC, close to 160 in CCP and 157 in UM591, but more than 106-140 in other plant pathogenic fungi (Figure 3, Additional file 1: Table S6). The 156 protein kinases of HGCC were classified into 8 groups. STE group (16 kinases) and MAPK family (5 kinases) in CMGC group were involved in MAPK pathway, and PKA family (3 kinases) of AGC group in cAMP pathway, CLK family (1 kinase) and RCK family (1 kinase) of CMGC group, CAMK group (22 kinase), and PKC family (2 kinases) in AGC group are related to Ca pathway. Interestingly, almost all protein kinases (151/156) of HGCC were PHI-associated genes, suggesting that protein kinases played key roles in pathogenic processes mainly mediated by the three pathogenic signal pathways. Fungal histidine kinase (HK) phosphorelay signaling pathway, i.e. two-component signaling pathway, play important roles in stress adaptation and virulence [32]. HGCC, CCP and UM591 contained 11, 13 and 12 HKs, respectively.

Protein families involved in degrading plant cuticle, cell wall, cytomembrane and cell inclusion

In order to successfully invade into host plant, plant pathogenic fungi are expected to produce and secrete multiple extracellular degrading enzymes such as cutinase for the degradation of cuticle [33], cell wall degrading enzymes (pectinase, cellulase and hemicellulase) [34], and cytomembrane and cell inclusion degrading enzymes (protease and lipase) [35, 36]. 9 cutinases were identified both in HGCC, CCP and UM591 genomes (Table S1). But the number was constricted (P = 0.02717) with other plant fungi as the reference (average 12). Pectate lyase, a type of pectinase, well existed in *C. cassiicola* (36 in HGCC, 31 in CCP, 34 in UM591), far more than 9-20 in other fungi. But pectinesterase, another type of pectinase was less in *C. cassiicola* (4 in HGCC, 4 in CCP, 5 in UM591) and other fungi (1-5). 17, 19 and 19 cellulase were identified in HGCC, CCP and UM591, respectively, close to the average (15) of plant pathogenic fungi. Glucanase, glucosidase and xylanase belonging to hemicellulase were rich in *C. cassiicola*, of which only glucosidase existed gene expansion (P = 0.00277). In HGCC, CCP and UM591 genomes, proteases were the second largest family including 73 subfamilies, but they were mainly in families of metallpo peptidase (138/144/136) and serine protease (372/376/374) (Figure 3, Additional file 1: Table S7). Aspartic peptidases are virulence factors both in plant and mammalian pathogens due to their ability in cleaving a large number of host proteins [37]. Interestingly, the number of A11 transposon peptidase (11 in HGCC, 10 in CCP, 8 in UM591) in *C. cassiicola* was significantly expanded (P = 0) in comparison with the average

(0.3) of other plant fungi. HGCC, CCP and UM591 contained 23, 27 and 24 lipases, respectively, which were not significantly expanded (P = 0.09042) compared to other plant fungi.

Glycoside hydrolase (GH) catalyzes the hydrolysis of glycosidic bonds in complex sugars [38], which functions as fungal pathogenesis. *C. cassiicola* contains 44 GH families (Additional file 1: Table S8). The numbers of GHs in HGCC (269) is lightly less than CCP (282) and UM591 (281), but far more than the average of other plant pathogenic fungi (207). More than half of HGCC GHs (143/269) were PHI-associated genes. It was reported that GH6, GH7, GH45 and GH61 cellulases and GH10 xlyanases were absent in insect pathogenic fungi, but present in plant pathogenic fungi [39]. The GH families of cellulases well existed in HGCC (72), CCP (74) and UM591 (75) including GH3, GH 6, GH7, GH45 and GH61 cellulases, of which GH3 and GH61 of cellulases were the majority. GH16 family of xyloglucosy transferases play an important role in the digestion of plant cell walls, and were well present in HGCC (14), CCP (17) and UM591 (18), close to the average (16) of other plant fungi.

Protein families for transportation

C. cassiicola contained a large number of transporters (693 in HGCC, 690 in CCP, 686 in UM591), which were classed into 88 families (Additional file 1: Table S9). The major facilitator superfamily (MFS) (240 in HGCC, 241 in CCP and 237 in UM591) and ABC (51 in HGCC, 50 in CCP and 48 in UM591) superfamily were the two biggest superfamilies. The former was capable of transporting small solutes in response to chemiosmotic ion gradients, and the latter transported small molecules and macromolecules under ATP hydrolysis [40, 41]. Notably, almost all MFS transporters and all ABC transporters in HGCC were PHI-associated genes, showing that MFS and ABC transporters played key roles in fungal pathogenicity.

In phytopathogenic fungi, drug transporters of ABC and MFS superfamilies can secrete endogenous fungal virulence factors such as toxin and protect pathogen against exogenous plant defense compounds such as phytoalexins [39, 42]. Two drug:H^+ antiporter (DHA) subfamilies (DHA1 and DHA2), drug transporters of MFS superfamily, can secrete toxic compounds into outer environment [43]. The multidrug resistance (MDR) and the pleiotropic drug resistance (PDR) subfamilies are drug transporters of ABC superfamily, being capable of functioning in resisting antifungal agents [43]. There was almost no difference in the numbers of DHA1 (41, 41 and 41), DHA2 (5, 6 and 6), MDR (10, 9 and 9) and PDR (15, 15 and 14) among HGCC, CCP and UM591 genomes (Additional file 1: Table S10), but they were more than the average of other plant pathogenic fungi. Interestingly, all drug transporters of HGCC were PHI-associated genes except for DHA2 (4/5).

Protein families for detoxification

Cytochrome P450 enzymes (CYPs) are proteins of a superfamily being ubiquitous in all biological kingdoms , which contain heme as a cofactor and therefore are hemoproteins. Fungal CYPs play key roles in various metabolisms such as housekeeping biochemical reactions, detoxification of chemicals and adaptation to adverse surroundings [44]. A large number of CYPs were identified in *C. cassiicola* HGCC (229), CCP (226) and UM591 (216), being classed into 115 families, which were far more than 95-167 in

other plant fungi (Figure 3, Additional file 1: Table S11). Interestingly, almost all CYPs of HGCC (211/229) were involved in the PHI. CYP65 and CYP505 subfamilies participate in the biosynthesis of mycotoxin, for example, CYP65 subfamily catalyzed the epoxidation reaction in the trichothecene biosynthesis in *F. graminearum* [45], and both CYP505 and CYP65 were required in the fumonisin biosynthesis of *Fusarium verticillioides* [46, 47], showing that CYP505 and CYP65 subfamilies were probably related to the mycotoxin biosynthesis in *C. cassiicola*. CYP65 was the biggest subfamily in CYP superfamily of *C. cassiicola* (24 in HGCC, 28 in CCP and 26 in UM591) and other plant pathogenic fungi. Relatively, *C. cassiicola* contained less CYP505 (6 in HGCC, 5 in CCP and 9 in UM591).

Secondary metabolite backbone genes

Melanin and mycotoxin are important virulence factors in plant pathogenic fungi [48]. So far, melanin and mycotoxin associated genes were not identified in *C. cassiicola* yet except for cassiicolin coded *Cas* gene. Secondary metabolite backbone genes were essential for the biosynthesis of melanin and mycotoxin as secondary metabolites. *C. cassiicola* HGCC contained 52 backbone genes, close to 51 in CCP and 49 in UM591 (Additional file 1: Table S12). Notably, almost all backbones (50/52) were PHI-associated genes, showing that these backbone genes were probably involved in the pathogenic process of *C. cassiicola*. The 52 backbone genes were classified into 5 groups including non-ribosomal peptide synthetase (NRPS), NRPS-like, polyketone synthase (PKS), PKS-like, dimethylallyl tryptophan synthase (DMAT) but lacking hybrid PKS-NRPS enzyme (HYBRID), of which PKS was the biggest group containing 33 genes. The numbers of backbone genes (P = 0.00350) and PKS (0.00032) were significantly expanded compared to other plant fungi.

In order to analyze the relationship between PKS domain and its function, phylogenetic analysis for the ketoacyl CoA synthase (KS) domain of PKS was performed among *C. cassiicola* PKSs and role-known PKSs in other pathogenic fungi (Figure 4). These PKSs were divided into two different clusters based on the phylogenetic analysis. One kind was reducing PKSs with KS, acyltransferase (AT) and dehydratase (DH) domains at least, which contained 25 *C. cassiicola* PKSs and 6 known PKSs involving in the biosynthesis of mycotoxin in other plant pathogenic fungi such as *Aspergillus ochraceus* AoLC35-12 for ochratoxin, *Alternaria alternate* ACTTS3 for ACT-toxin, *B. maydis* PKS1 and PKS2 for T-toxin, *Gibberella zeae* PKS4 for zearalenon, and *Gibberella moniliformis* Fum1p for fumonisin. Two reducing PKSs (g13578 and g7009) were specific in CCPagainst HGCC. The other kind was non-reducing PKSs with KS and AT domains at least and without dehydratase (DH), enoyl reductase (ER) and ketoreductase (KR) domains, which included 10 PKSs of *C. cassiicola* and 9 known PKSs related to melanin biosynthesis in other fungal pathogens such as *Aspergillus fumigatus* Alb1p, *Ceratocystis resinifera* PKS1, *Colletotrichum lagenarium* PKS1, *Chaetomium globosum* PKS-1, *Ascochyta rabie* PKS1, *B. maydis* PKS18, *S. turcica* StPKS, *Bipolaris oryzae* PKS1, and *A. alternata* ALM1. It was hard to identify which reducing PKSs were involved in the mycotoxin biosynthesis of *C. cassiicola* due to the complicated evolutionary relationship for KS domain of *C. cassiicola* reducing PKSs and known mycotoxin related PKSs. Nevertheless, a non-reducing PKS HGCC_7666 of *C. cassiicola* had the closest evolutionary relationship with known melanin-

associated non-reducing PKSs, suggesting that HGCC_7666 was probably the backbone gene for melanin synthesis of *C. cassiicola*.

Small, cysteine-rich peptides and effector proteins

The small cysteine-rich proteins (SCRPs) could be secreted directly into host plant cells to function as host recognition or colonization [49] and the stimulation of host hypersensitive response (HR) [50]. Some SCRPs as virulence effectors facilitated fungal virulence by multiple ways including perturbing host cell signaling, interfering with host recognition of the pathogen and suppressing pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI)[51]. Some SCRPs as *Avr* genes triggered or suppressed effector-triggered immunity (ETI) mediated by a gene-for-gene system in the PHI [52]. 51 SCRPs were identified in HGCC ranging in size from 68 to 150 amino acids, including 3 hydrophobin and 1 cerato-platanin (Additional file 1: Table S13). Fungal hydrophobins are involved in surface recognition [53], and play a role as effectors in the PHI [54]. Cerato-platanin family proteins could elicit disease resistance responses in the host plant [55].

Lysin motif (LysM) contained effectors are ubiquitous in plant pathogenic fungi, which are secreted out of the cell to suppress the immune response of host [56, 57]. Common in fungal extracellular membrane (CFEM) domain is a fungi specific domain containing eight cysteines and is found in some proteins with proposed roles in fungal infection and colonization [58, 59]. HGCC contained 8 LysM-contained and 20 CEEM-contained genes, of which 5 LysM-contained and 15 CFEM-contained genes had signal peptides. It was suggested that the 5 LysM and 15 CFEM as candidate effectors were probably involved in fungal pathogenic process.

Gene expression in spore germination

To reveal the molecular and cellular processes during spore germination of *C. cassiicola*, 6 h- and 12 h-germination times were selected for RNA-seq based transcriptome analysis. For 12 h at 25℃ in sterile ddH_2O, the majority of spores were germinated at two ends, and only a few of that at one end (Figure 1B). Germinated and un-germinated spores of *C. cassiicola* were harvested in biological triplicates for RNA isolation. RNA-seq libraries were constructed and sequenced using Illumina Hiseq™. Following quality control and adapters trimming, 611,310,902 bp clean paired reads were obtained from 6 *C. cassiicola* RNA-seq libraries. Sequencing yield from individual libraries ranged from 84,671,942 to 116,800,420 reads per sample (Table S14 in Additional file 2). Pearson correlation analysis show that samples of same treatment were high correlative to each other in gene expression levels with 0.885-1 of R^2 value higher than 0.742-0.85 of R^2 value between samples in different treatments (Figure S1), which suggested that samples in the same treatment were repeatable and RNA-seq data was reliable. 66-73% of *C. cassiicola* filtered reads per library mapped to the gene predictions of *C. cassiicola*. A total of 3,288 genes were differentially expressed, of which 1,552 and 1,736 were up-regulated and down-regulated during spore germiantion of *C. cassiicola* with un-germinated spores as the reference, respectively (Figure 5). Specific

DEGs were listed in Table S15 in Additional file 2. These DEGs contained hundreds of previously identified functional gene from the genome sequences classified by gene families (Table 3).

A total of 2,600 DEGs (79.08%) were annotated by gene ontology (GO). GO term analysis for *C. cassiicola* revealed an enrichment for cellular process, localization, metabolic process, regulation of biological process and single-organism process, and binding and catalytic activity (Figure 6). A total of 891 DEGs (34.27%) were annotated by KEGG pathway, containing 624 up-regulated DEGs and 267 down-regulated DEGs. The pathway classification of DEGs were listed in Table S16 in Additional file 2, and the top 20 of the enriched pathway terms were shown in Figure 7. The pathway enrichment statics indicated that the majority of the KEGG annotated DEGs were involved in metabolism, genetic information processing, cellular processes, organismal system, human diseases and environmental information processing.

qRT-PCR validation of selected DEGs

qRT-PCR assays were conducted to validate the gene expression patterns of 66 DEGs containing 49 up-regulated and 17 down-regulated genes. As shown in Figure 8, qRT-PCR data were correlated with the RNA-seq data (R^2=0.7565). The express patterns of 56 DEGs were confirmed by qRT-PCR except for 10 DEGs (Table S17). These results showed a high correlation of RNA-seq and qRT-PCR results, indicating that the RNA-seq data was very reliable.

Discussion

C. cassiicola is a fungal phytopathogen, capable of infecting hundreds of plants and causing leaf spot disease [1]. In this study, we reported the genome sequence and the spore germination-related transcriptome of the cucumber-sampled *C. cassiicola* isolate. Genome sequencing of the pathogen provided large amounts of important genetic information for us, which extremely contributed to understand its evolutionary relationship with other species of pathogenic fungi and screen pathogenicity-associated genes. A great lot of DEGs related to spore germination were identified through RNA-seq. These results would facilitate the study of pathogenic mechanism of *C. cassiicola* to cucumber.

Phylogenomic analysis showed that *C. cassiicola* had more close genetic affinity with other Pleosporales order fungi than not Pleosporales order fungi, which were similar to previously reported results [24]. Therefore, it was convenient and efficient to dig novel functional factors of *C. cassiicola* being involved in the PHI and study the pathogenic mechanism of *C. cassiicola* based on the comparative genomics with other well researched Pleosporales order fungi.

As a phytopathogenic fungus, *C. cassiicola* needs undergo several crucial and complicated steps to cause host plant disease such as attachment, germination, infection structures formation, invasion and colonization [60]. Fungal pathogen could successfully complete these processes directly via a great number of virulence factors including cutinases, cell wall degrading enzymes, cytomembrane and cell inclusion degrading enzymes, mycotoxin, melanin and effectors, *et al*. The production, transportation and regulation of these virulence factors are mediated by multiple pathways of genes including MAPK, Ca and

cAMP pathogenic signal pathways of genes, transcription factors, transporters, core genes for the biosynthesis of secondary metabolite, CYPs, *et al* [13]. Even more, there are cross-talks among multiple pathways, which forms an extremely complicated regulation network. Therefore, the high throughput identification of PHI-associated genes would accelerate the study for pathogenic mechanism of *C. cassiicola*. As expected, thousands of putative virulence-associated genes in various families or pathways were identified in *C. cassiicola* HGCC, which provided insights into the pathogenic mechanism of *C. cassiicola* to cucumber. Although HGCC, CCP and UM591 were conspecific and close to each other in relationship revealed by phylogenomic analysis, there were differences in sequence features of genome among HGCC, CCP and UM591 isolates sampled from different plants. The phenomenon could be explained that the three isolates evolved and developed specific genetic material to adapt adverse surroundings. Therefore, the specific information was worthy of being played more attention.

Fungal infection to plants usually requires contact of conidia with host and subsequently germination [20]. Conidial germination is a genetically programmed and highly coordinated phenomenon, which involves the initiation of biochemical activity increase in metabolism and the induction of morphological changes [61, 62]. Isotropic growth, also named as swelling, is the first morphological change, which is accompanied by water uptake, cell wall growth, cellular composition changes and cytoplasmic micro-viscosity decrease [63], and by many metabolic activities. After swelling, chitin in cell wall becomes polarized, and the fungal cell extends at a restricted area at the tip of the cell, which results in the elongation of germ tube [64, 65].

Gene expression profiles can provide insight into molecular and cellular processes of spore germination either in transcriptional level or in proteome level. Zhou *et al*. found that a total of 3,026 gene were differentially expressed during spore germination of *Penicillium expansum* using RNA-seq and iTRAQ, and most of them were involved in metabolism and genetic information processing [20]. Liu *et al.* identified 66 DEGs during spore germination of *Nosema bombycis* using RNA-seq [21]. Joise Hander *et al.* used two-dimensional SDS-PAGE and mass spectrometry to identify a total of 316 spore germination associated proteins in *Moniliophthora perniciosa*, including fungal filamentation associated proteins Septin and Kinesin, a fumagillin associated transcription factor and polyketide synthase, ATP synthase, binding immunoglobulin protein (Bip), and catalase [66]. Bassi *et al.* found 1,646 DEGs containing toxin-associated genes *nheC*, *cytK* and *hblC* during *Bacillus thuringiensis* spore germination [67]. These research results showed that the global transcriptional and protein level analysis were effective tools to understand the molecular and cellular processes of spore germination. Therefore, RNA-seq was used to explore global gene expression during spore germination of *C. cassiicola* in this study. A total of 3,288 genes were found to be differentially expressed, which was similar to 3,026 DEGs during spore germination of *P. expansum* [20]. It was found from GO analysis results of *C. cassiicola* and *P. expansum* that the majority of spore germination associated DEGs were enriched in cellular process, localization, metabolic process, regulation of biological process, single-organism process, binding, and catalytic activity. Similar pathway classification results for DEGs were also found between *C. cassiicola* and *P. expansum* that the majority of KEGG-annotated DEGs were classified in metabolism, Genetic information processing, environmental information processing, cellular processes, and human diseases. This

phenomenon showed that RNA-seq data for spore germination was reliable, which probably provide valuable information for revealing the molecular mechanism of *C. cassiicola* germination.

Carbohydrate metabolism involves the various biochemical processes responsible for the formation, breakdown and interconversion of carbohydrates, which serve as short-term fuel for organisms [20]. Therefore, it is not surprising during a relatively short time of germination that 152 DEGs were involved in carbohydrate metabolism, including glycolysis / gluconeogenesis (29 DEGs), citrate cycle (12 DEGs), pentose phosphate pathway (15 DEGs), pentose and glucuronate interconversions (16 DEGs), fructose and mannose metabolism (12 DEGs), galactose metabolism (13 DEGs), ascorbate and aldarate metabolism (9 DEGs), starch and sucrose metabolism (17 DEGs), amino sugar and nucleotide sugar metabolism (24 DEGs), pyruvate metabolism (26 DEGs), glyoxylate and dicarboxylate metabolism (21 DEGs), propanoate metabolism (13 DEGs), butanoate metabolism (17 DEGs), C5-Branched dibasic acid metabolism (3 DEGs), and Inositol phosphate metabolism (5 DEGs). Most of these DEGs were up-regulated, showing that carbohydrate metabolism was essential for the spore germination of *C. cassiicola*. The similar results were found in spore germination process of *Penicillium expansum* [20]. Energy metabolism can convert biochemical energy into adenosine triphosphate (ATP). 73 gene were differentially expressed during the spore germination of *C. cassiicola*, and most of these DEGs were up-regulated, which was nearly adverse to zhou's results that all DEGs being involved in energy metabolism were down-regulated both in transcriptional and protein levels.

Conclusions

In conclusion, the genome sequence of *C. cassiicola* from cucumber was presented, and thousands of virulence associated genes were mined in genome-wide by homologous search against multiple functional databases and conserved domain searches. 3,288 genes were differentially expressed during the spore germination of *C. cassiicola*. Most of KEGG annotated DEGs were involved in metabolism, genetic information processing, cellular processes, organismal system, human diseases and environmental information processing. These results would not only facilitate to understand the molecular pathogenic mechanism to cucumber and the molecular and cellular processes during the spore germination of *C. cassiicola*, but also laid the foundation for the disease control.

Methods

Fungal strain, hyphae collection, and DNA isolation

C. cassiicola (Berk. & M.A. Curtis) strain HGCC, isolated from infected cucumber leaves by us in Shanghai in 2010, was highly virulent to cucumber, thus it was used for genome sequence. The isolate was cultured on potato dextrose agar (PDA) medium in test tubes at 25℃ for 7 d and then maintained at 4℃. HGCC isolate was sub-cultured on PDA plates at 25℃ in dark for 10 d, and then its mycelium was collected through peeling off using a glass slide, which was ground in liquid nitrogen. Genomic DNA was extracted

from finely grounded material using CTAB method [23]. DNA quality control was ensured by 1% agarose gel electrophoresis and Infinite M200 PRO (Tecan, Switzerland).

Spore sampling, RNA extraction and sequencing

Spore production test was conducted on PDA plates. After being activated on PDA plates at 25℃, HGCC strain was sub-cultured on PDA plates at 25℃ for 10 d with 12 h light and 12 h darkness per day. 5 mL of sterile ddH_2O was added on each plate and spores were peeled off using sterile writing brushes and filtered with three layers of sterile gauzes to prepare spore suspension. Spores were collected through centrifugation with 7,000 rpm for 5 min. Spores were re-suspended and adjusted to 10^5 spores/mL of spore suspension with sterile ddH_2O. The spore suspension was incubated at 25℃ with darkness for 6 h and 12 h to perform spore germination test. Equal volumes of 6 h and 12 h germinated spore suspension were mixed. The germinated spores in the mixed suspension were collected by centrifugation with 7000 rpm for 5 min. The germinated spores and un-germinated spores were grinded for 1 min using a Geno/Grinder 2010 (SPEX SamplePrep, USA) after being freezed rapidly with liquid nitrogen for the extraction of total RNA. Total RNA were extracted using TaKaRa MiniBEST Universal RNA Extraction Kit following the manufacturer's recommended method and saved at -80℃. The test was conducted in three biological duplication.

Total RNA was quantified using a Nanodrop and assessed for purity using a Nanodrop and for integrity using a Agilent 2100. RNA samples returning a RNA integrity number (RIN) value greater than 6.3 were considered acceptable for sequencing. A cDNA library was made from each qualified RNA sample. The libraries were sequenced using Illumina Hiseq^{TM} at Novogene Bioinformatics Institute, Beijing.

Whole genome shotgun sequencing and assembly

The whole genome of HGCC was de novo sequenced using Illumina Hiseq X-Ten in a 2×149 bp paired-end mode with a 400 bp insert sizes of library at Personalbio (Shanghai, China). High quality data was obtained from raw data through removing adapter contamination with AdapterRemoval (version 2) [68], collection with SOAPec (version 2.01) [69], and length screening with a threshold (> 50 bp). These highly quality data were de novo assembled using SPAdes v3.9.0 [70].

Gene prediction, annotation, identification of gene orthology analyses, and interspecific phylogenomic analysis

Gene structures of *C. cassiicola* genome sequences were predicted using Augustus software with the annotated gene information of *B. cinerea* and HGCC transcripts as references [71]. Putative secreted proteins were identified by combining Target 1.1 [72], SignalP 4.1 Server [73], TMHMM Server v. 2.0 [74], and Big-GPI softwares [75]. Putative small cystein-rich proteins (SCRPs) were screened from secreted proteins based on their sequence characteristics such as 20-150 amino acids and at least four cysteins [76]. Protein family classification of *C. cassiicola* were performed by sequence alignment against Pfam

database with profile hidden Markov models (http://xfam.org/) using hmmer 3.1b2 (http://www.hmmer.org/).

Potential virulence-associated genes were identified by local Blastp searches of *C. cassiicola* HGCC protein sequences against PHI database (version 4.4, http://www.phi-base.org/) with a cutoff *E*-value of 1e-5. MAPK pathway associated genes of *C. cassiicola* were screened by local Blastp against *S. cerevisiae* MAPK pathway associated genes with a cutoff *E*-value of 1e-5 that were confirmed by experiments. GPCRs were identified by local Blastp against GPCRDB database (http://gpcrdb.org/) with best hits and further confirmed by searching seven transmembrane helices with TMHMM Server v. 2.0. Kinases, proteases and transporters were identified by local Blastp against the KinBase database with a cutoff *E*-value of 1e-10 and the MEROPS peptidase database with a cutoff *E*-value of 1e-20, Transporter Classification Database (TCDB) [77] with a cutoff *E*-value of 1e-40, respectively [78]. Cytochrome P450s and glucoside hydrolases (GHs) families were classified based on Blastp alignment against P450 database (http://drnelson.uthsc.edu/CytochromeP450.html) and CAZy database (http://www.cazy.org/) with a cutoff *E*-value of 1e-10, respectively.

Secondary metabolism backbone genes such as PKS, NRPS and NRPS-PKS were identified in SMURF system (http://smurf.jcvi.org/index.php) [79]. *C. cassiicola* PKSs and other functionally known PKSs in other fungi were submitted to SBSPKS database (http://www.nii.ac.in/~pksdb/sbspks/master.html) to modulate and analyze conserved domains. Phylogenetic analysis of PKSs was performed by aligning of ketoacyl CoA synthase (KS) domain sequences and creating a Maximum Likelihood tree using MEGA 7.0 with Jones-Taylor-Thornton (JTT) model. These functionally known PKSs related toxin biosynthesis included *A. ochraceus* AoLC35-12 (GenBank: AAT92023) for ochratoxin, *A. alternate* ACTTS3 (GenBank: BAJ14522) for ACT-toxin, *B. maydis* PKS1 (GenBank: AAB08104) and PKS2 (GenBank: AAR90257) for T-toxin, *G. zeae* PKS4 (GenBank: ABB90283) for the biosynthesis of zearalenones and *G. moniliformis* Fum1p (GenBank: AAD43562) for fumonisin. Melanin-associated PKSs included *A. fumigatus* Alb1p (GenBank: ACJ13039), *Ceratocystis resinifera* PKS1 (GenBank: AAO60166), *C. lagenarium* PKS1 (GenBank: BAA18956), *C. globosum* PKS-1 (GenBank: AFP82905), *A. rabie* PKS1 (GenBank: ACS74449), *B. maydis* PKS18 (GenBank: AAR90272), *S. turcica* StPKS (GenBank: AEE68981), *B. oryzae* PKS1 (GenBank: BAD22832), *A. alternate* ALM1 (GenBank: BAK64048).

To constructed an interspecific phylogeny tree among *C. cassiicola* and other 10 fungal species including *P. nodorum*, *P. tritici-repentis*, *S. turcica*, *C. zeae-maydis*, *B. maydis*, *C. lunata*, *A. flavus*, *B. cinerea*, *M. oryzae*, *F. graminearum*. 2,344 core protein sequences were screened by local reciprocal Blast search with a cutoff of *E*-value 1e-20 and more than 80% sequence identity. These orthologous proteins were aligned using Clustal W 2.1 [80]. A maximum Likelihood tree was created by the concatenated amino acid sequences using MEGA7.0 with JTT model.

Gene expression in spore germination in *C. cassiicola*

The quality of the raw reads generated from RNA-sequencing were checked with FastQC [81]. The clean reads were obtained from the raw reads by removing adapter containing reads, >10% “N” containing

reads, and reads of low quality. The 6 trimmed *C. cassiicola* RNA-seq libraries were mapped on the predicted CDS of *C. cassiicola* using Bowtie2 with default settings [82]. The number of reads mapped to each gene for each RNA set was calculated from the .sam alignment files derived from Bowtie2.

The counts of RNA-seq reads over transcripts were used to calculate the fold change of gene expression using DESeq2 [83]. DEGs were selected by the cutoffs both on padj and on log2FoldChange. Genes were considered differentially expressed if the padj<0.05 and |log2FoldChange|>1. Pathway enrichments of DEGs were performed by Kyoto Encyclopaedia of Genes and Genomes (KEGG).

Quantitative real-time PCR (qRT-PCR) assays

66 DEGs involving in cellular processes were selected for the confirmation of the RNA-seq data using qRT-PCR (Table S16 in Additional file 2). *Gapdh* gene was used as the internal control. The first strand of cDNA was synthesized from total RNA of germinated and un-germinated spores using EasyScript All-in-One First-Strand cDNA Synthesis SuperMix for qPCR (One-Step gDNA Removal) (Transgen, China). The qRT-PCR test was conducted on an QuanStudio 6 Flex (Thermo Flasher Scientific, USA). 20 μL of reaction mixture contained 2 μL diluted cDNA (1-100 ng), 0.4 μL each primer (10 μM), 7.2 μL nuclease-free water and 10 μL 2×TransStart® Top Green qPCR SuperMix (Transgen, China). The reaction program was follows: one cycle at 94℃ for 30 s, and 40 cycles at 94℃ for 5 s, 60℃ for 15 s and 72℃ for 10 s. The specificities of the amplification products were checked through a melting-curve analysis between 60℃ and 95℃ after each PCR reaction. The relative gene expression data was analyzed using $2^{-\Delta\Delta Ct}$ method [84]. These qRT-PCR assays were performed with three biological and three technical replicates.

Abbreviations

MAPK: mitogen-activated protein kinase; GPCR: G-protein-coupled receptor; PHI: pathogen-host interaction; GH: glycoside hydrolase; ABC: ATP-binding cassette; MFS: major facilitator superfamily; DHA: drug:H+ antiporter; MDR: multidrug resistance; PDR: pleiotropic drug resistance; CYP: cytochrome P450 enzyme; NRPS: non-ribosomal peptide synthetase; PKS: polyketone synthase; HK: histidine kinase; DMAT: dimethylallyl tryptophan synthase; HYBRID: hybrid PKS-NRPS enzyme; KS: ketoacyl CoA synthase; AT: acyltransferase; DH: dehydratase; ER: enoyl reductase; KR: ketoreductase; SCRP: small cysteine-rich protein; HR: hypersensitive response; PAMP: pathogen-associated molecular pattern; PTI: PAMP-triggered immunity; ETI: effector-triggered immunity; LysM: lysin motif; CFEM: common in fungal extracellular membrane; KEGG: Kyoto Encyclopaedia of Genes and Genomes; DEG: Differentially expressed gene; CDS: Coding sequence.

Declarations

Acknowledgements

Not applicable.

Funding

This work was supported by Natural Science Foundation of Shanghai (Grant No. 16ZR1424100), SAAS Program for Excellent Research Team (Grant No. 2017(A-03)), China National Major Program of Science and Technology (Grant No. 2017ZX07202004-004), and Shanghai Agriculture Applied Technology Development Program, China (Grant No. G2014070204).

Availability of data and materials

The data that support the findings of this study are available from [https://www.ncbi.nlm.nih.gov/genome/, Accession No. RJLO00000000] but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of [https://www.ncbi.nlm.nih.gov/genome/, Accession No. RJLO00000000].

The dataset(s) supporting the conclusions of this article is(are) included within the article (and its additional file(s)).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing financial and non-financial interests.

Authors' contributions

SG conducted the submission of genome data, protein family classification, the identification of functional genes, and drafted the manuscript. RZ carried out DNA and total RNA extraction, sequencing associated works, and phylogenetic analysis. YS analyzed the RNA-seq data. LX took charge qRT-PCR. ZS and PG revised the manuscript. FD conceived the study and designed the article contents. All authors had read and approved the final manuscript.

Authors' information

Shigang Gao, Email: shggao@163.com

Rong Zeng, Email: superzero@126.com

Yujuan Suo, Email: yujsuo@163.com

Lihui Xu, Email: xulihui@saas.sh.cn

Zhiwei Song, Email: songzhiwei061@163.com

Ping Gao, Email: 26065574@qq.com

Fuming Dai, Email: fumingdai@163.com

References

1. Li BJ, Gao W, Shi YX, Xie XW. Progress in researches on Corynespora leaf spot. Acta Phytophy Sin. 2012; 39(2):171-176.
2. Sumabat L, Kemerait RC, Brewer MT. Phylogenetic diversity and host specialization of *Corynespora cassiicola* responsible for emerging target spot disease of cotton and other crops in the southeastern United States. Phytopathology. 2018; 108(7):892-901.
3. Chung SR, Lee HS, You IC. Fungal keratitis caused by *Corynespora cassiicola*, a plant pathogen. J Mycol Infect. 2018; 23(1):24-26.
4. Xie ZL, Wu W, Meng DS, Zhang Q, Ma YQ, Liu W, Chen JH. A case of Phaeohyphomycosis caused by *Corynespora cassiicola* BMC Infect Dis. 2018; 18(1):444.
5. Liu MT, Zhang DF, Sun HT. Studies on the biological characteristics of *Corynespora cassiicola*. China Vegetables. 2003; 20(2):93-102.
6. Liu D, Qin Z, Zhang Y, Zhou X, Xin M. Histological observation of cucumber infected with *Corynespora cassiicola*. Eur J Plant Pathol. 2017; 149(2):455-466.
7. Dixon LJ, Schlub RL, Pernezny K, Datnoff LE. Host specialization and phylogenetic diversity of *Corynespora cassiicola*. Phytopathology. 2009; 99(9):1015.
8. Shimomoto Y, Sato T, Hojo H, Morita Y, Takeuchi S, Mizumoto H, Kiba A, Hikichi Y. Pathogenic and genetic variation among isolates of *Corynespora cassiicola* in Japan. Plant Pathol. 2015; 60(2):253-260.
9. Barthe P, Pujaderenaud V, Breton F, Gargani D, Thai R, Roumestand C, De FL. Structural analysis of cassiicolin, a host-selective protein toxin from *Corynespora cassiicola*. J Mol Biol. 2007; 367(1):89-101.
10. Déon M, Fumanal B, Gimenez S, Bieysse D, Oliveira RR, Shuib SS, Breton F, Elumalai S, Vida JB, Seguin M. Diversity of the cassiicolin gene in *Corynespora cassiicola* and relation with the pathogenicity in *Hevea brasiliensis*. Fungal Biol-UK. 2014; 118(1):32-47.
11. Wu J, Xie X, Shi Y, Chai A, Wang Q, Li B. Erratum to: Variation of cassiicolin genes among Chinese isolates of *Corynespora cassiicola*. J Microbiol. 2018; 56(9):691.
12. Liu ZH, Ye QI, Huang XY, Yang H, Hou Y, Zhang R. Conditions and activity analysis of cell wall degrading enzymes produced from *Corynespora cassiicola* of brown spot of cucumber. China Vegetables.

13. Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, Thon M, Kulkarni R, Xu JR, Pan H *et al*. The genome sequence of the rice blast fungus *Magnaporthe grisea*. Nature. 2005; 434(7036):980-986.
14. Liu XM, Pu JJ, Zhang X, Qi YX, Xie YX, Zhang H, Zheng FC. Cloning and bioinformatics analysis of gene *CCK1* related to pathogenicity of *Corynespora cassiicola* on Hevea Rubber. Biotech Bull. 2012; 0(10):168-172.
15. Qi YX, Zhang X, Pu JJ, Lu Y, Zhang H, Zhang HQ, Xie YX. Cloning and sequence analysis of a Slt2-type MAPK homologous gene *CMP1* from *Corynespora cassiicola* of *Hevea brasiliensis*. Chinese J Trop Crop. 2010; 31(11):1951-1958.
16. Liu XM, Zheng FC, Zhang X, Qi YX, Xie YX, Zhang H, Pu JJ. Cloning and pathogenicity of toxin gene *cc004-cas* from *Corynespora cassiicola* on *Hevea brasiliensis*. Chinese J Trop Crop. 2014; 35(9):1809-1815.
17. Déon M, Bourré Y, Gimenez S, Berger A, Bieysse D, De LF, Poncet J, Roussel V, Bonnot F, Oliver G. Characterization of a cassiicolin-encoding gene from *Corynespora cassiicola*, pathogen of rubber tree (*Hevea brasiliensis*). Plant Sci. 2012; 185-186(4):227-237.
18. Leeuwen MRV, Krijgsheld P, Bleichrodt R, Menke H, Stam H, Stark J, Wösten HAB, Dijksterhuis J. Germination of conidia of *Aspergillus niger* is accompanied by major changes in RNA profiles. Stud Mycol. 2013; 74(1):59-70.
19. Dague E, Alsteens D, Latgé JP, Dufrêne YF. High-resolution cell surface dynamics of germinating *Aspergillus fumigatus* Biophys J. 2008; 94(2):656-660.
20. Zhou T, Wang X, Luo J, Ye B, Zhou Y, Zhou L, Lai T. Identification of differentially expressed genes involved in spore germination of *Penicillium expansum* by comparative transcriptome and proteome approaches. Microbiologyopen. 2018; 7(3):e00562.
21. Liu H, Li M, He X, Cai S, He X, Lu X. Transcriptome sequencing and characterization of ungerminated and germinated spores of *Nosema bombycis*. Acta Biochim Biophys Sin. 2016; 48(3):246-256.
22. Nagler K, Krawczyk AO, De Jong A, Madela K, Hoffmann T, Laue M, Kuipers OP, Bremer E, Moeller R. Identification of differentially expressed genes during *Bacillus subtilis* spore outgrowth in high-salinity environments using RNA Sequencing. Front Microbiol. 2016; 7:1564.
23. Gao SG, Li YQ, Gao JX, Suo YJ, Fu KH, Li YY, Chen J. Genome sequence and virulence variation-related transcriptome profiles of *Curvularia lunata*, an important maize pathogenic fungus. Bmc Genomics. 2014; 15(1):627.
24. Lopez D, Ribeiro S, Label P, Fumanal B, Venisse JS, Kohler A, De RO, Labutti K, Lipzen A, Lail K. Genome-wide analysis of *Corynespora cassiicola* leaf fall disease putative effectors. Front Microbiol. 2018; 9:276.
25. Hong KL, Yue FT, Su MY, Ling S, Kuan CS. Genomic insight into pathogenicity of dematiaceous fungus *Corynespora cassiicola*. Peerj. 2017; 5(19):e2841.
26. Affeldt KJ, Carrig J, Amare M, Keller NP. Global survey of canonical *Aspergillus flavus* G protein-coupled receptors. Mbio. 2014; 5(5):01501-01514.

27. Li L, Wright SJ, Krystofova S, Park G, Borkovich KA. Heterotrimeric G protein signaling in filamentous fungi. Annu Rev Microbiol. 2007; 61(1):423.
28. DeZwaan TM, Carroll AM, Valent B, Sweigard JA. *Magnaporthe grisea* pth11p is a novel plasma membrane protein that mediates appressorium differentiation in response to inductive substrate cues. Plant Cell. 1999; 11(10):2013-2030.
29. Preininger AM, Hamm HE. G protein signaling: insights from new structures. Sci STKE. 2004(218):re3.
30. Svoboda P, Teisinger J, Novotny J, Bourova L, Drmota T, Hejnova L, Moravcova Z, Lisy V, Rudajev V, Stohr J *et al*. Biochemistry of transmembrane signaling mediated by trimeric G proteins. Physiol Res. 2004; 53 Suppl 1:S141-152.
31. Yu HY, Seo JA, Kim JE, Han KH, Shim WB, Yun SH, Lee YW. Functional analyses of heterotrimeric G protein G alpha and G beta subunits in *Gibberella zeae*. Microbiol. 2008; 154(Pt 2):392-401.
32. Rispail N, Di PA. The two-component histidine kinase Fhk1 controls stress adaptation and virulence of *Fusarium oxysporum*. Mol Plant Pathol. 2010; 11(3):395-407.
33. Gui Y, Zhang W, Zhang D, Zhou L, Short D, Wang J, Ma X, Li T, Kong Z, Wang BL. A verticillium dahliae extracellular cutinase modulates plant immune responses. Mol Plant Microbe Interact. 2017; 31(2):260-273.
34. Castellmiller CV, Gutierrezgonzalez JJ, Tu ZJ, Bushley KE, Hainaut M, Henrissat B, Samac DA. Genome assembly of the fungus *Cochliobolus miyabeanus*, and transcriptome analysis during early stages of infection on American Wildrice (*Zizania palustris*). Plos One. 2016; 11(6):e0154122.
35. Gouran H, Gillespie H, Nascimento R, Chakraborty S, Zaini PA, Jacobson A, Phinney BS, Dolan D, Durbinjohnson BP, Antonova ES. The secreted protease PrtA controls cell growth, biofilm formation and pathogenicity in *Xylella fastidiosa*. Scientific Reports. 2016; 6:31098.
36. Gaillardin C: Lipases as pathogenicity factors of fungi. In: Timmis KN, editors. Handbook of hydrocarbon and lipid microbiology. Springer, Berlin, Heidelberg; 2010. p. 3259-3268.
37. Coetzer THT, Goldring JPD, Huson LEJ. Oligopeptidase B: A processing peptidase involved in pathogenesis. Biochimie. 2008; 90(2):336-344.
38. Bourne Y, Henrissat B. Glycoside hydrolases and glycosyltransferases: families and functional modules. Curr Opin Struc Biol. 2001; 11(5):593-600.
39. Zheng P, Xia YL, Xiao GH, Xiong CH, Hu X, Zhang SW, Zheng HJ, Huang Y, Zhou Y, Wang SY *et al*. Genome sequence of the insect pathogenic fungus *Cordyceps militaris*, a valued traditional Chinese medicine. Genome Biol. 2011; 12(11):R116.
40. Dimroth P. Primary sodium ion translocating enzymes. Biochimica et Biophysica Acta (BBA) - Bioenergetics. 1997; 1318(1–2):11-51.
41. Paulsen IT, Brown MH, Skurray RA. Proton-dependent multidrug efflux systems. Microbiol Rev. 1996; 60(4):575-608.
42. de Waard MA. Significance of ABC transporters in fungicide sensitivity and resistance. Pestic Sci. 1997; 51(3):271-275.

43. Roohparvar R, De Waard MA, Kema GH, Zwiers LH. MgMfs1, a major facilitator superfamily transporter from the fungal wheat pathogen *Mycosphaerella graminicola*, is a strong protectant against natural toxic compounds and fungicides. Fungal Genet Biol. 2007; 44(5):378-388.
44. Durairaj P, Hur JS, Yun H. Versatile biocatalysis of fungal cytochrome P450 monooxygenases. Microbial Cell Factories. 2016; 15(1):1-16.
45. Kimura M, Tokai T, Takahashi-Ando N, Ohsato S, Fujimura M. Molecular and genetic studies of fusarium trichothecene biosynthesis: pathways, genes, and evolution. Biosci Biotechnol Biochem. 2007; 71(9):2105-2123.
46. Butchko RA, Plattner RD, Proctor RH. Deletion analysis of FUM genes involved in tricarballylic ester formation during fumonisin biosynthesis. J Agric Food Chem. 2006; 54(25):9398-9404.
47. Bojja RS, Cerny RL, Proctor RH, Du LC. Determining the biosynthetic sequence in the early steps of the fumonisin pathway by use of three gene-disruption mutants of *Fusarium verticillioides*. J Agr Food Chem. 2004; 52(10):2855-2860.
48. Gao JX, Chen J. Transcriptome analysis identifies candidate genes associated with melanin and toxin biosynthesis and pathogenicity of the maize pathogen, *Curvularia lunata*. J Phytopathol. 2018; 166(4):233-241.
49. Rep M, Van der Does HC, Meijer M, Van Wijk R, Houterman PM, Dekker HL, De Koster CG, Cornelissen BJC. A small, cysteine-rich protein secreted by *Fusarium oxysporum* during colonization of xylem vessels is required for *I-3*-mediated resistance in tomato. Mol Microbiol. 2010; 53(5):1373-1383.
50. Sornkom W, Miki S, Takeuchi S, Abe A, Asano K, Sone T. Fluorescent reporter analysis revealed the timing and localization of AVR-Pia expression, an avirulence effector of *Magnaporthe oryzae*. Mol Plant Pathol. 2017; 18(8):1138.
51. Zheng X, Wagener N, Mclellan H, Boevink PC, Hua C, Birch PRJ, Brunner F. *Phytophthora infestans* RXLR effector SFI5 requires association with calmodulin for PTI/MTI suppressing activity. New Phytol. 2018; 219(4):1433-1446.
52. Medina CA, Reyes PA, Trujillo CA, Gonzalez JL, Bejarano DA, Montenegro NA, Jacobs JM, Joe A, Restrepo S, Alfano JR. The role of type III effectors from *Xanthomonas axonopodis* manihotis in virulence and suppression of plant immunity. Mol Plant Pathol. 2018; 19(3):593-606.
53. Talbot NJ, Kershaw MJ, Wakley GE, De Vries O, Wessels J, Hamer JE. MPG1 encodes a fungal hydrophobin involved in surface interactions during infection-related development of *Magnaporthe grisea*. Plant Cell. 1996; 8(6):985.
54. Casarrubia S, Daghino S, Kohler A, Morin E, Khouja HR, Daguerre Y, Veneaultfourrey C, Martin FM, Perotto S, Martino E. The hydrophobin-like OmSSP1 may be an effector in the ericoid mycorrhizal symbiosis. Front Plant Sci. 2018; 9:546.
55. Frías M, Brito N, González C. The *Botrytis cinerea* cerato-platanin BcSpl1 is a potent inducer of systemic acquired resistance (SAR) in tobacco and generates a wave of salicylic acid expanding from the site of application. Mol Plant Pathol. 2013; 14(2):191-196.

56. Mentlak TA, Kombrink A, Shinya T, Ryder LS, Otomo I, Saitoh H, Terauchi R, Nishizawa Y, Shibuya N, Thomma BP *et al*. Effector-mediated suppression of chitin-triggered immunity by *Magnaporthe oryzae* is necessary for rice blast disease. Plant Cell. 2012; 24(1):322-335.
57. 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(5994):953-955.
58. Kulkarni RD, Kelkar HS, Dean RA. An eight-cysteine-containing CFEM domain unique to a group of fungal membrane proteins. Trends Biochem Sci. 2003; 28(3):118-121.
59. Zhang ZN, Wu QY, Zhang GZ, Zhu YY, Murphy RW, Liu Z, Zou CG. Systematic analyses reveal uniqueness and origin of the CFEM domain in fungi. Sci Rep-UK. 2015; 5:13032.
60. Schafer W. Molecular mechanisms of fungal pathogenicity to plants. Annu Rev Phytopathol. 1994; 32(1):461-477.
61. Moir A. How do spores germinate? J Appl Microbiol. 2006; 101(3):526-530.
62. Osherov N, May G. Conidial germination in *Aspergillus nidulans* requires RAS signaling and protein synthesis. Genetics. 2000; 155(2):647-656.
63. Van Leeuwen MR, Van Doorn TM, Golovina EA, Stark J, Dijksterhuis J. Water- and air-distributed conidia differ in sterol content and cytoplasmic microviscosity. Appl Environ Microbiol. 2010; 76(1):366-369.
64. Taheri-Talesh N, Horio T, Araujo-Bazan L, Dou X, Espeso EA, Penalva MA, Osmani SA, Oakley BR. The tip growth apparatus of Aspergillus nidulans. Mol Biol Cell. 2008; 19(4):1439-1449.
65. Leeuwen MRV, Smant W, Boer WD, Dijksterhuis J. Filipin is a reliable in situ marker of ergosterol in the plasma membrane of germinating conidia (spores) of *Penicillium discolor* and stains intensively at the site of germ tube formation. J Microbiol Meth. 2008; 74(2):64-73.
66. Mares JH, Gramacho KP, Santos EC, da Silva Santiago A, Santana JO, de Sousa AO, Alvim FC, Pirovani CP. Proteomic analysis during of spore germination of *Moniliophthora perniciosa*, the causal agent of witches' broom disease in cacao. BMC Microbiol. 2017; 17(1):176.
67. Bassi D, Colla F, Gazzola S, Puglisi E, Delledonne M, Cocconcelli PS. Transcriptome analysis of *Bacillus thuringiensis* spore life, germination and cell outgrowth in a vegetable-based food model. Food Microbiol. 2016; 55:73-85.
68. Schubert M, Lindgreen S, Orlando L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. Bmc Research Notes. 2016; 9(1):88.
69. Luo RB, Liu BH, Xie YL, Li ZY, Huang WH, Yuan JY, He GZ, Chen YX, Pan Q, Liu YJ. SOAPdenovo2: an empirically improved memory-efficient short-readde novoassembler. GigaScience. 2012; 1(1):18-18.
70. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012; 19(5):455-477.
71. Stanke M, Tzvetkova A, Morgenstern B. AUGUSTUS at EGASP: using EST, protein and genomic alignments for improved gene prediction in the human genome. Genome Biol. 2006; 7 Suppl 1:S11

11-18.
72. Nielsen H, Engelbrecht J, Brunak S, von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. 1997; 10(1):1-6.
73. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat methods. 2011; 8(10):785-786.
74. Chen Y, Yu P, Luo J, Jiang Y. Secreted protein prediction system combining CJ-SPHMM, TMHMM, and PSORT. Mamm Genome. 2003; 14(12):859-865.
75. Eisenhaber B, Bork P, Eisenhaber F. Sequence properties of GPI-anchored proteins near the omega-site: constraints for the polypeptide binding site of the putative transamidase. Protein Eng. 1998; 11(12):1155-1161.
76. Graham MA, Silverstein KAT, Cannon SB, VandenBosch KA. Computational identification and characterization of novel genes from legumes. Plant Physiol. 2004; 135(3):1179-1197.
77. Saier MH, Tran CV, Barabote R. TCDB: the Transporter Classification Database for membrane transport protein analyses and information. Nucleic Acids Res. 2006; 34(Database issue):181-186.
78. Schomburg D, Schomburg I. Enzyme databases. Methods Mol Biol. 2010; 609:113-128.
79. Khaldi N, Seifuddin FT, Turner G, Haft D, Nierman WC, Wolfe KH, Fedorova ND. SMURF: Genomic mapping of fungal secondary metabolite clusters. Fungal Genet Biol. 2010; 47(9):736-741.
80. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R *et al*. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23(21):2947-2948.
81. Andrews S. Babraham Bioinformatics: FastQC A Quality Control tool for High Throughput Sequence Data. In.; 2010. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
82. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012; 9(4):357-359.
83. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014; 15(12):550.
84. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 2001; 29(9):e45.

Tables

Table 1 Comparison of genome features among three *C. cassiicola* isolates HGCC, CCP and UM591.

Features	HGCC	CCP	UM591
Assembly size (Mb)	42.7	44.8	41.4
Scaffolds	1032	244	1941
GC (%)	51.78	51.89	52.47
Repeated sequences (%)	0.52	-	-
Protein-coding genes	15678	15614	15388
Gene density (genes per Mb)	367.2	348.5	371.7
Average CDS length (Bp)	1404	1410	1397
Secreted proteins	1166	1216	1182

Table 2 Genome-wide analysis for *C. cassiicola* HGCC, CCP and UM591 gene sets.

Characteristics	HGCC	CCP	UM591	core[a]	HG-CCP-UM591 core[b]	HG-CCP-UM591 specific[c]	HG specific[d]	CCP specific[e]	UM591 specific[f]	HGCC PHI
Protein-encoding genes	15678	15614	15388	12337	13672	1335	1047	961	971	4490
Secreted proteins	1166	1216	1182	994	1107	113	29	37	34	402
PHI	4490	4523	4415	4365	4370	5	100	109	82	-
Proteases	638	649	634	620	623	3	11	13	9	248
Glycoside hydrolases	269	282	281	263	263	0	6	7	3	143
Lipase	23	27	24	23	23	0	0	0	1	9
Glucanase	32	35	31	32	32	0	1	1	0	25
Glucosidase	13	13	16	12	12	0	1	0	0	13
Xylanase	7	8	8	7	7	0	0	0	0	4
Pectate lyase	36	31	34	35	35	0	1	0	0	30
Cutinase	9	9	9	9	9	0	0	0	0	6
Cellulase	17	19	19	17	17	0	0	0	0	4
Pectinesterase	4	4	5	4	4	0	0	0	0	4
Lipase	23	27	24	23	23	0	0	0	1	9
MFS transporters	240	241	237	239	239	0	1	1	1	231
ABC transporters	51	50	48	51	51	0	0	0	0	51
P450s	229	226	216	228	228	0	1	0	0	211
Protein kinases	156	160	157	153	153	0	3	5	5	151
Histidine kinase	11	13	12	11	11	0	0	0	0	11
GPCRs	181	165	169	176	178	2	2	1	6	102
Pth11-like GPCRs	77	61	55	75	75	0	1	1	1	62
CFEM-containing proteins	21	21	19	19	20	1	1	1	0	14
Fungal specific transcription factors	213	229	221	203	203	0	9	6	5	198
Backbone genes for secondary metabolism	52	51	49	49	49	0	3	3	0	50

[a]core, *C. cassiicola* HGCC, CCP, UM591 and other 10 phytopathogenic ascomycetes genes grouped by reciprocal blast analysis with a cutoff *E* value of 1e-5; [b]HGCC-CCP-UM591 core, HGCC genes both present in CCP and UM591grouped by reciprocal blast analysis with a cutoff *E* value of 1e-5; [c]HGCC-CCP-UM591 specific, HGCC genes both present in CCP and UM591 but specific in other 10 fungi grouped by reciprocal blast analysis with a cutoff

E value of 1e-5; [d]HG specific, specific genes of HGCC against CCP, UM591 and other 10 fungi; [e]CCP specific, specific genes of CCP against HGCC, UM591 and other 10 fungi; [f]UM591 specific, specific genes of UM591 against HGCC, CCP and other 10 fungi.

Table 3 Spore germination-related gene families.

Gene family	Gene number	Up-regulated gene number	Down-regulated gene number
GPCR	63	18	45
G-protein alpha subunit	1	1	0
Protein kinase	36	19	17
Glycoside hydrolase	76	42	34
P450	102	23	79
Protease	234	110	124
Pectate lyase	13	4	9
Cellulase	8	4	4
Cutinase	4	1	3
Glucanase	12	8	4
Glucosidase	4	4	0
Xylanase	2	1	1
SCRP	18	5	13
PKS	12	1	11
NRPS	3	0	3
NRPS_like	2	1	1
Transporter	243	94	149
ABC transporter	15	3	12
MFS transporter	94	33	61

Additional File Legends

Figure S1 Pearson correlation between samples in gene expression levels.

Additional file 1: Table S1. The number of genes for selected gene families in *C. cassiicola* and other ascomycetes.

Additional file 1: Table S2. G protein coupled receptors in different fungal genomes.

Additional file 1: Table S3. Gene-encoding proteins for MAPK pathway in *C. cassiicola* HGCC.

Additional file 1: Table S4. Gene-encoding proteins for cAMP signal pathway in *C. cassiicola* HGCC.

Additional file 1: Table S5. Gene-encoding proteins for Ca signal pathway in *C. cassiicola* HGCC.

Additional file 1: Table S6. The number of protein kinases in different fungal genomes.

Additional file 1: Table S7. Protease genes classed by MEROPS family in different fungal genomes.

Additional file 1: Table S8. Glycoside hydrolase in *C. cassiicola* and other phytopathogenic ascomycetes.

Additional file 1: Table S9. Transporters in *C. cassiicola* HGCC, CCP and UM591.

Additional file 1: Table S10. Drug transporters in *C. cassiicola* and other phytopathogenic ascomycetes.

Additional file 1: Table S11. Cytochrome P450 genes classed by CYP family in different fungal genomes.

Additional file 1: Table S12. Numbers of backbone-genes for the biosynthesis of secondary metabolites in different pathogenic fungi.

Additional file 1: Table S13. The putative small, cysteine-rich peptides encoded genes in *C. cassiicola* HGCC.

Additional file 2: Table S14. Sequencing yield from individual libraries per sample.

Additional file 2: Table S15. 3,288 DEGs during spore germination identified by RNA-seq.

Additional file 2: Table S16. KEGG pathway enrichment analysis for DEGs.

Additional file 2: Table S17. qRT-PCR for selected DEGs involving in cell growth and death.

Figures

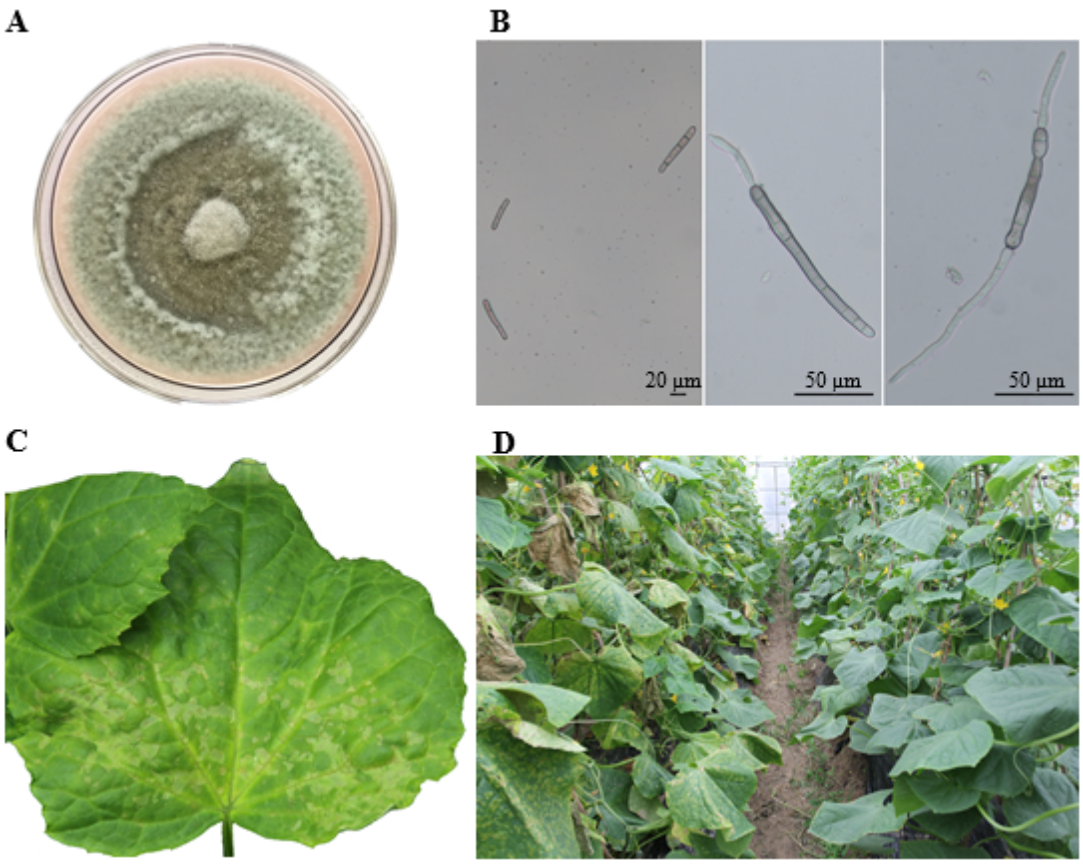

Figure 1

C. cassiicola isolate HGCC and cucumber target leaf spot symptoms. (A) HGCC mycelium colony on PDA medium 10 days after subculture at 25℃ in dark. (B) Optical microscopy observation of HGCC conidia with/no germination in water. (C) Cucumber target leaf spot symptoms on a susceptible cucumber (Shenqing-1) leaf inoculated by HGCC with 1×105 spores (60 h host-inoculation). (D) Cucumber target leaf spot symptoms on the susceptible cucumber Shenqing-1 (left), and no obvious symptoms on a resistant cucumber Biyu (right), in a greenhouse of Shanghai.

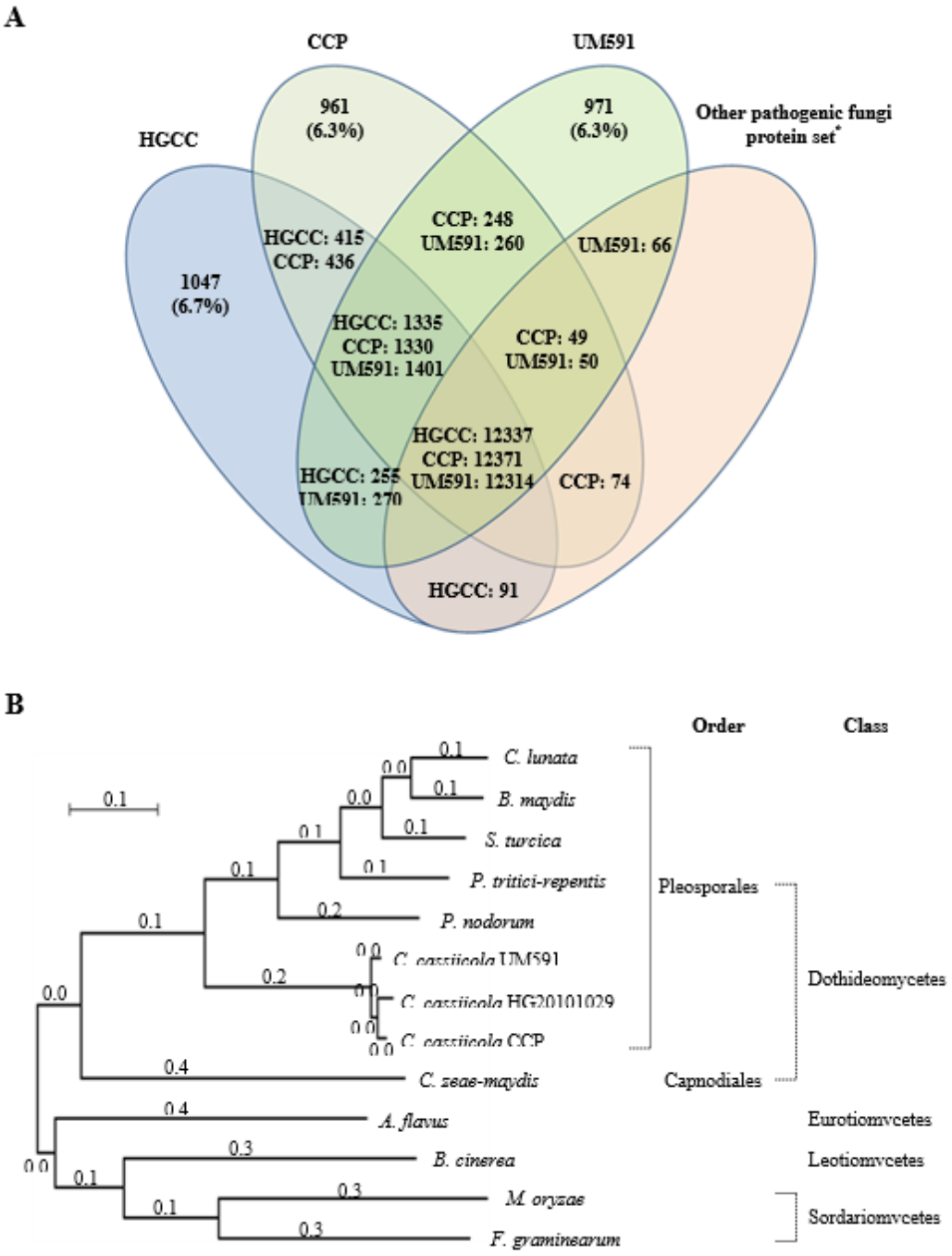

Figure 2

Comparative genomics and evolutionary analysis of C. cassiicola. (A) Reciprocal Blast analysis of the protein sequences among three C. cassiicola isolates and other plant pathogenic fungi. *Other pathogenic fungi protein set, protein set of select other plant pathogenic fungi used in evolutionary analysis. Evolutionary relationship of C. cassiicola with selected fungal species was shown by a maximum likelihood phylogenomic tree (B) constructed using MEGA 7.0 software.

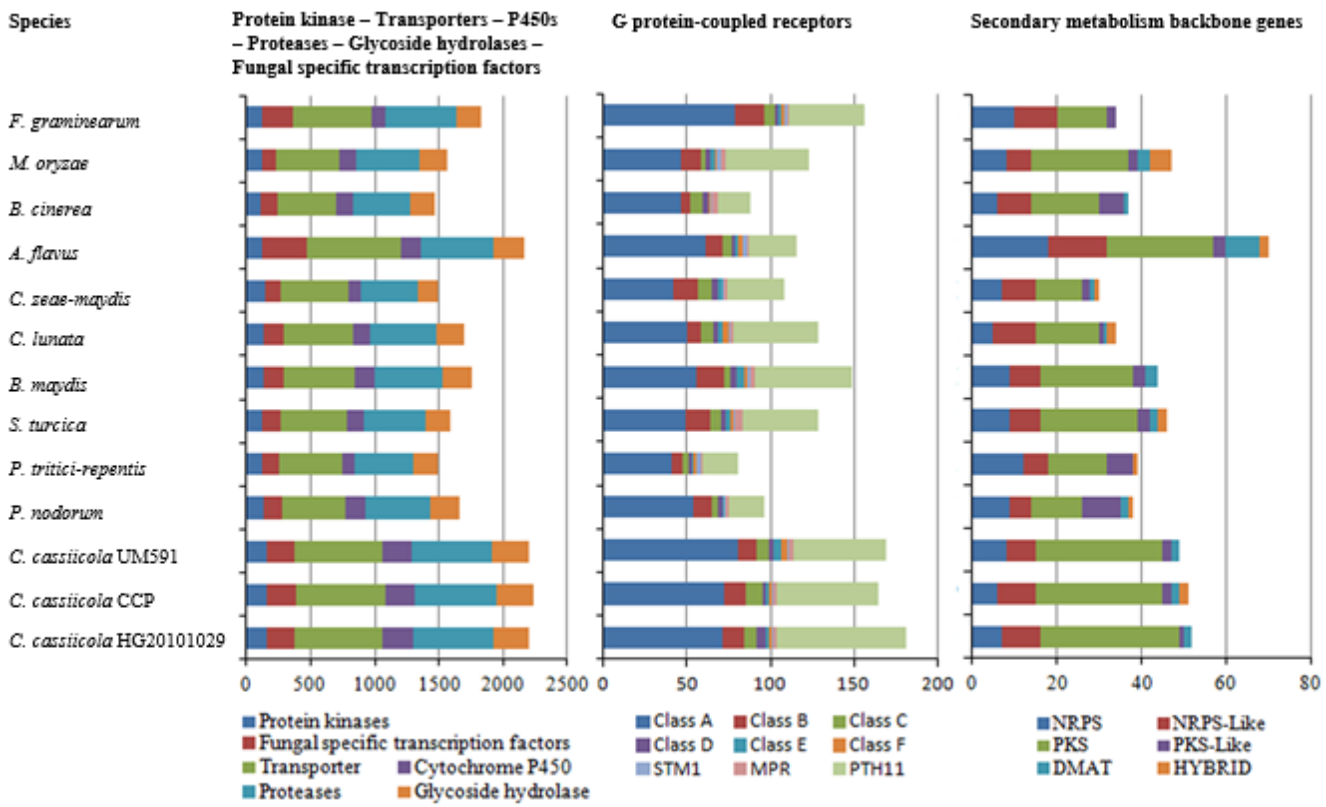

Figure 3

Virulence-associated protein family analysis of C. cassiicola and other fungi. NRPS, non-ribosomal peptide synthetase; PKS, polyketide synthetase; DMAT, dimethylallyl tryptophan synthase; HYBRID, hybrid PKS-NRPS enzyme.

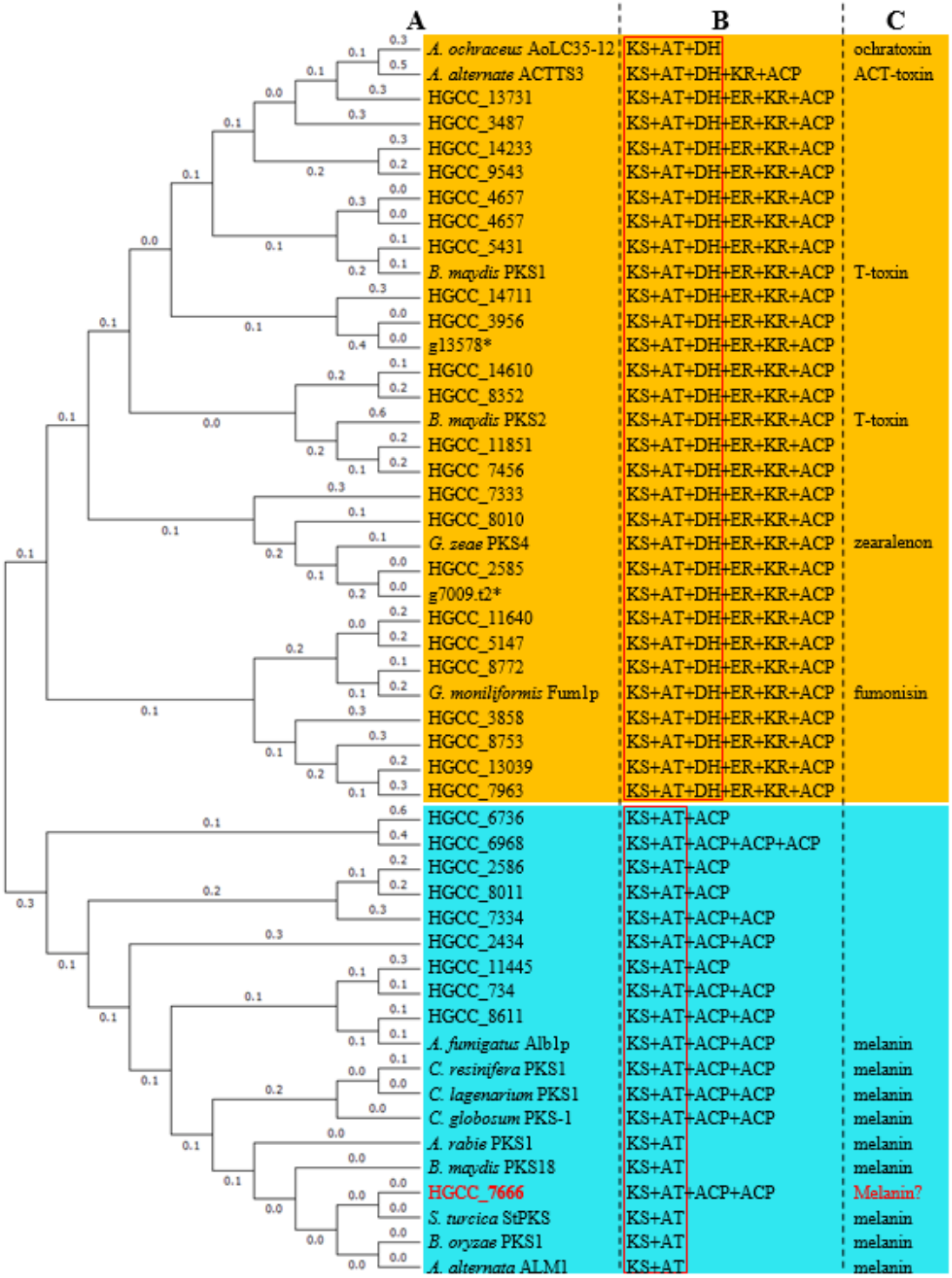

Figure 4

Phylogenetic and domain analysis of polyketide synthases (PKS) in C. cassiicola HGCC and other fungi. (A) A maximum likelihood tree of ketoacyl CoA synthase (KS) domain sequences of PKS in selected fungi. (B) Domain analysis of these PKS using SBSPKS database. Domain definitions: KS, ketoacyl CoA synthase; AT, acyltransferase domain; DH, dehydratase domain; ER, enoyl reductase domain; KR, ketoreductase domain; ACP, acyl carrier protein domain. (C) Toxin and melanin related to reported PKS in

other fungi, the accession No. of which are shown in the “Methods” section. *g13578 and *g7009 are specific in CCP against HGCC.

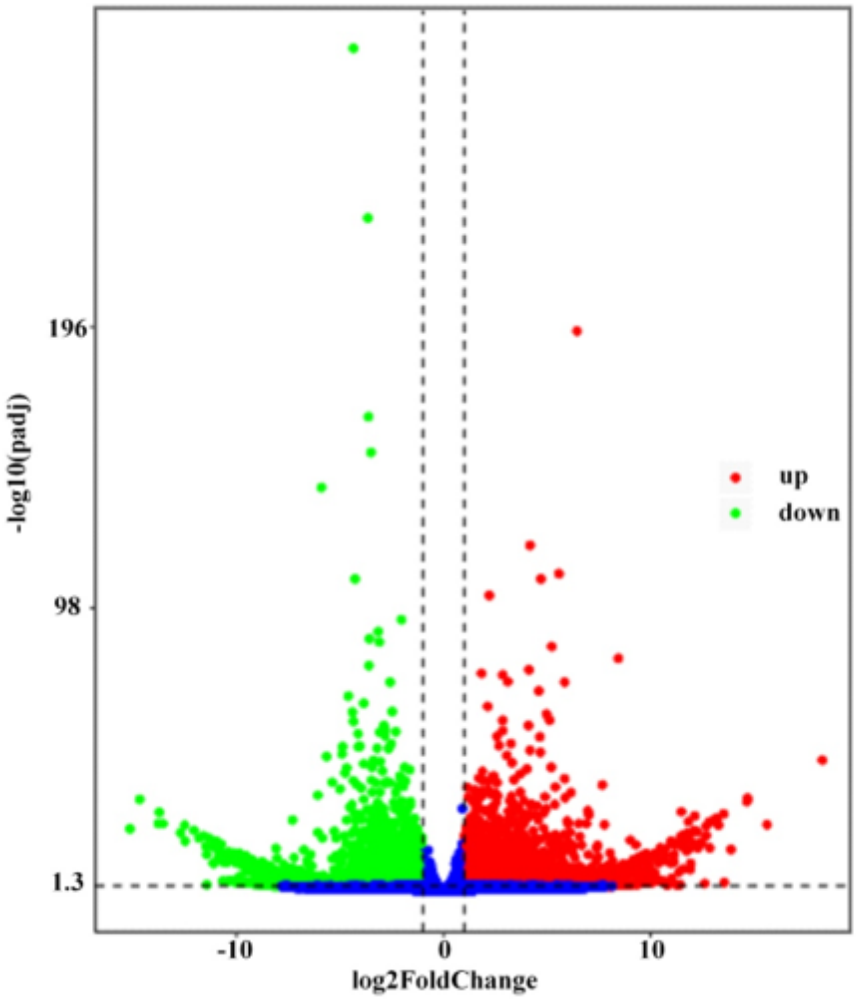

Figure 5

Volcano plot for global comparison of transcript profiles between germinated and ungerminated spores. Spots with red color mean up-regulated genes with a threshold of log2FoldChange>=1 and padj<0.05. Spot with green color mean down-regulated genes with a threshold of log2FoldChange >=-1 and padj <0.05.

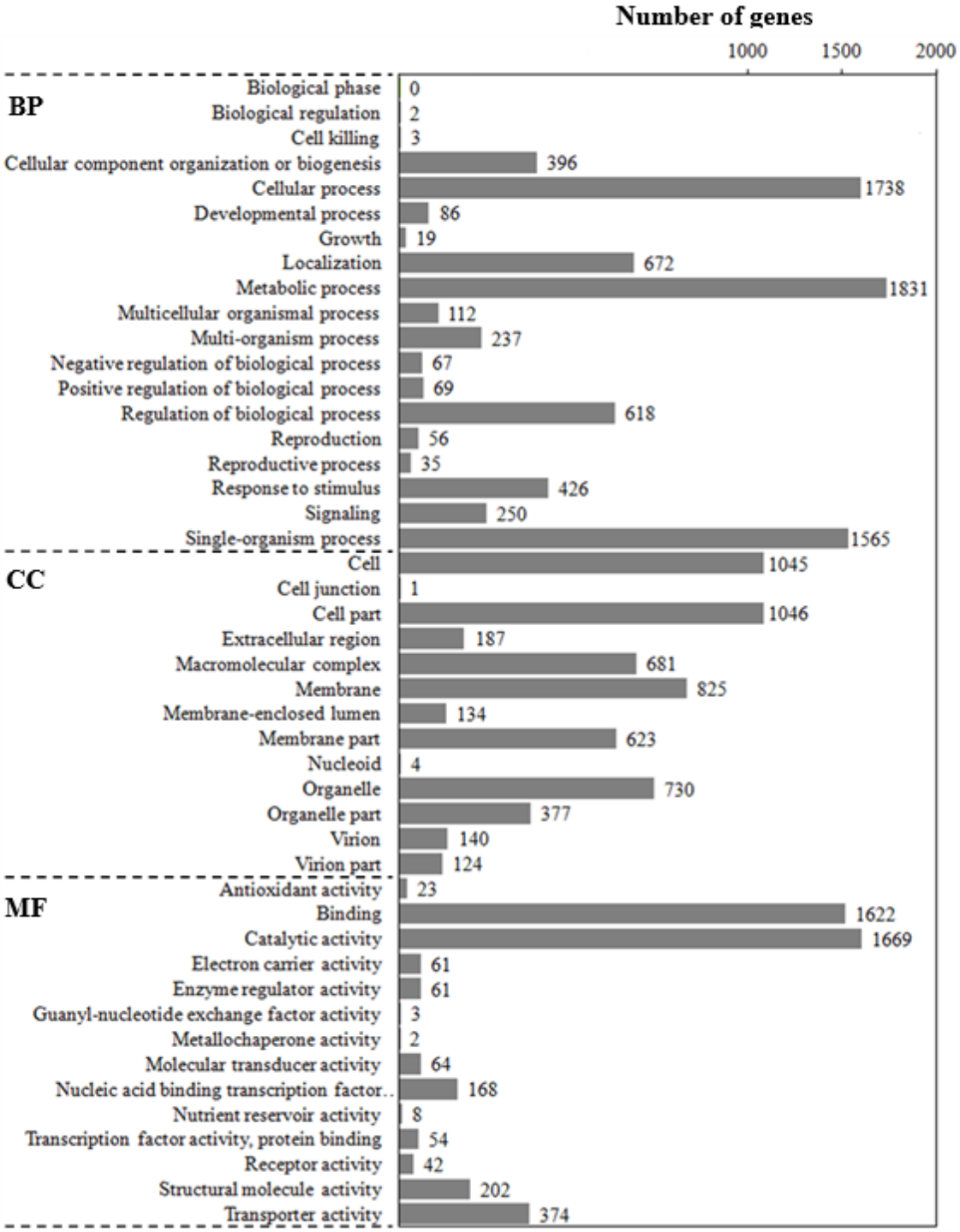

Figure 6

GO function classification of DEGs. BP: Biological process; CC: Cellular component; MF: Molecular function.

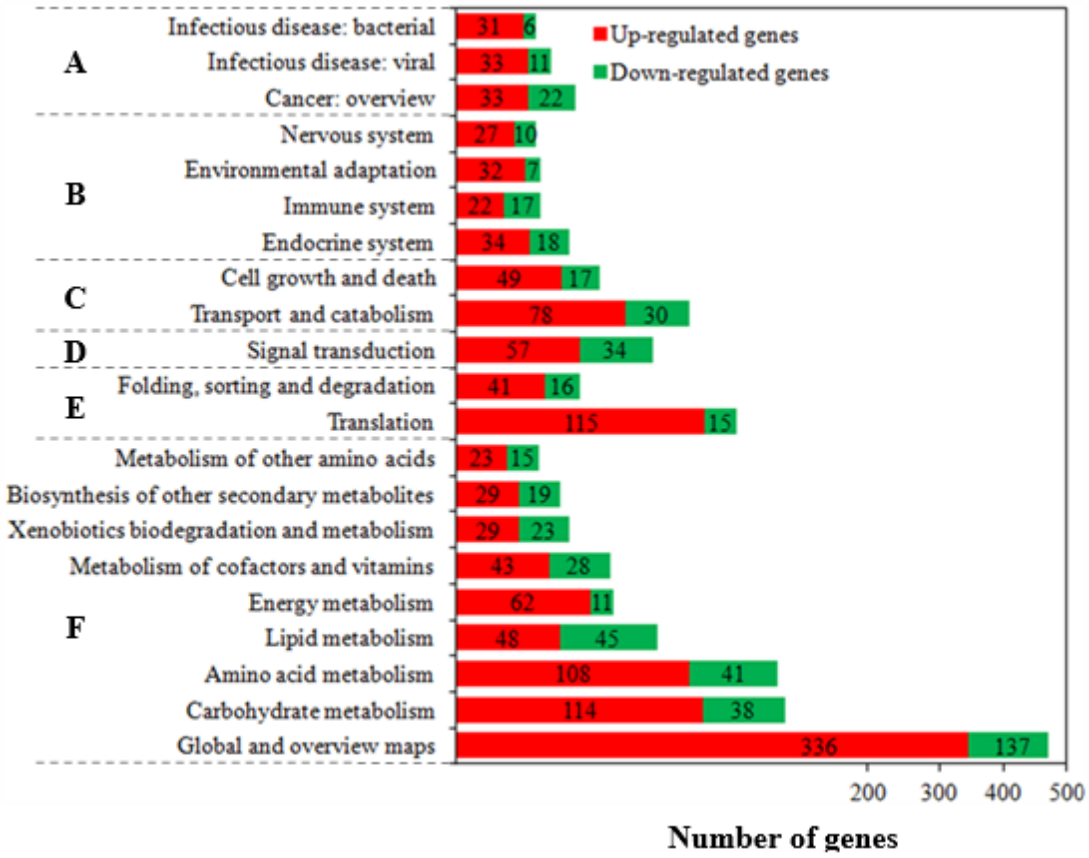

Figure 7

KEGG classification of DEGs. The top 20 of the second pathway terms are displayed. A, B, C, D, E and F mean the first pathway terms. A: Human diseases; B: Organismal systems; C: Cellular processes; D: Environmental information processing; E: Genetic information processing; F: Metabolism.

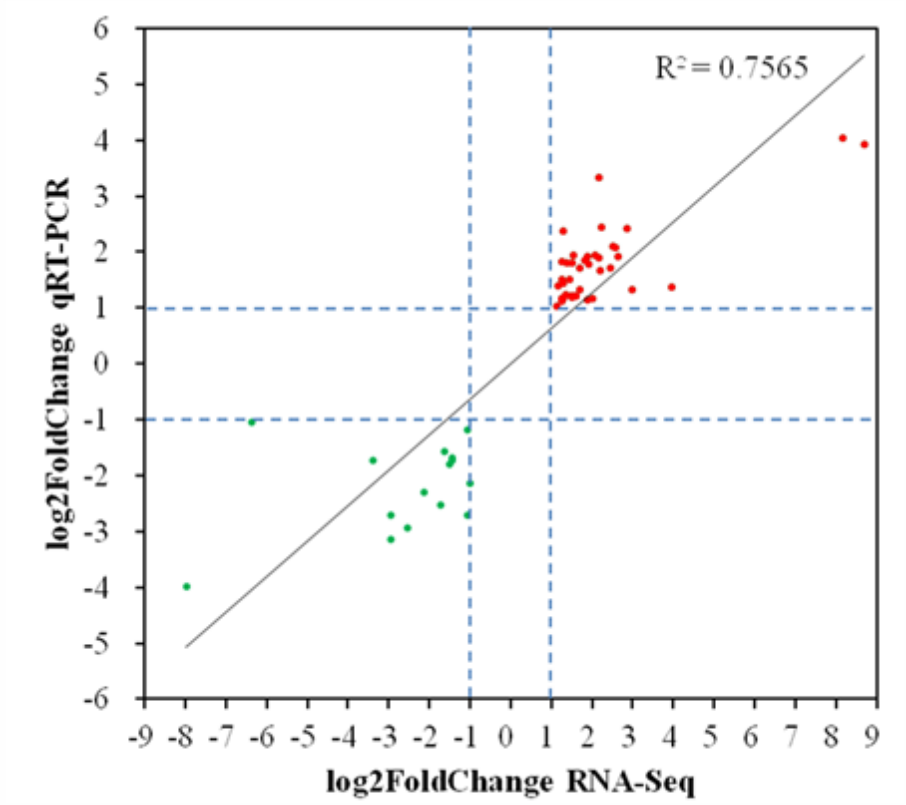

Figure 8

The correlation of RNA-Seq and qRT-PCR analysis for DEGs.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.xls
- FigureS1.pdf
- Additionalfile2.xlsx