Article

Not peer-reviewed version

Revealing Hidden Genes in *Botrytis cinerea*: New Insights into Genes Involved in the Biosynthesis of Secondary Metabolites

Ivonne Suárez , Isidro G. Collado * , Carlos Garrido *

Posted Date: 10 April 2024

doi: 10.20944/preprints202404.0687.v1

Keywords: *Botrytis cinerea*; secondary metabolism; polyketide synthase; non-ribosomal peptide synthetase; sesquiterpene cyclase; diterpene cyclase; dimethylallyltryptophan synthase; bioinformatics; fungal pathog

 doi:10.20944/preprints202404.0687.v1

Article

Revealing Hidden Genes in *Botrytis cinerea*: New Insights into Genes Involved in the Biosynthesis of Secondary Metabolites

Ivonne Suárez [1,2,3], Isidro G. Collado [2,3,*] and Carlos Garrido [1,4,*]

[1] Laboratorio de Microbiología, Departamento de Biomedicina, Biotecnología y Salud Pública, Facultad de Ciencias del Mar y Ambientales, Universidad de Cádiz, 11510, Puerto Real, Spain. (I.S. ivonne.suarez@uca.es)
[2] Departamento de Química Orgánica, Facultad de Ciencias, Campus Universitario Río San Pedro s/n, Torre sur, 4ª planta, Universidad de Cádiz, 11510, Puerto Real, Cádiz, Spain.
[3] Instituto de Investigación en Biomoléculas (INBIO), Universidad de Cádiz, 11510 Puerto Real, Cádiz, Spain.
[4] Instituto de Investigación Vitivinícola y Agroalimentaria (IVAGRO), Universidad de Cádiz, 11510 Puerto Real, Cádiz, Spain.
* Correspondence: isidro.gonzalez@uca.es (I.G.C.); carlos.garrido@uca.es (C.G.)

Abstract: Utilizing bioinformatics tools, this study expands our understanding of secondary metabolism in *Botrytis cinerea*, identifying novel genes within polyketide synthase (PKS), non-ribosomal peptide synthetase (NRPS), sesquiterpene cyclase (STC), diterpene cyclase (DTC), and dimethylallyltryptophan synthase (DMATS) families. These findings enrich the genetic framework associated with *B. cinerea's* pathogenicity and ecological adaptation, offering insights into uncharted metabolic pathways. Significantly, the discovery of previously unannotated genes provides new molecular targets for developing targeted antifungal strategies, promising to enhance crop protection and advance our understanding of fungal biochemistry. This research not only broadens the scope of known secondary metabolites but also opens avenues for future exploration into *B. cinerea's* biosynthetic capabilities, potentially leading to novel antifungal compounds. Our work underscores the importance of integrating bioinformatics and genomics for fungal research, paving the way for sustainable agricultural practices by pinpointing precise molecular interventions against *B. cinerea*. This study sets a foundation for further investigations into the fungus's secondary metabolism, with implications for biotechnology and crop disease management.

Keywords: *Botrytis cinerea*; secondary metabolism; polyketide synthase; non-ribosomal peptide synthetase; sesquiterpene cyclase; diterpene cyclase; dimethylallyltryptophan synthase; bioinformatics; fungal pathogenicity; antifungal targets

1. Introduction

Bioinformatics is an emerging biological discipline, constituting a unique blend of basic biology, genetics, and information sciences. It focuses on the technological advancements in the collection and processing of biological data, including DNA sequence data, RNA, and cDNA sequence data, as well as protein sequence information, and the utilization of this data for biological exploration and prediction using digital technology [1,2]. Over the past decades, a plethora of bioinformatics software and databases have been developed to analyze the flood of omics data (genomics, transcriptomics, proteomics, and metabolomics) and interpret their biological significance [3]. Due to this omics era, sequencing data are publicly available in databases such as GenBank, Ensembl and others (Table 1), which integrate nucleotide sequences and biological annotations, making them widely accessible for the identification of genes responsible for secondary metabolism [2].

 doi:10.20944/preprints202404.0687.v1

Table 1. Open access databases information for bioinformatics analysis [2].

Database	Scope of inclusion	Common application areas
GenBank	Contains 19.6 trillion base pairs from over 2.9 billion nucleotide sequences of 504,000 formally described species, collecting annotations of all publicly available nucleic acid and protein sequences. https://www.ncbi.nlm.nih.gov/ [4]	
Ensembl	High quality genomic resources for vertebrates and model organisms, including gene, variant, regulatory region and comparative genomic resources, annotation of repetitive sequences, genome-wide alignment, variant and regulatory characterization. https://www.ensembl.org [5]	Genomics, transcriptomics, proteomics, metabolomics
KEGG	It contains more than 8400 genomes, systematically analyzes gene functions, links genomic information and functional information, and includes 16 databases such as metabolic pathway database, hierarchical classification database, gene database, and genome database. https://www.kegg.jp [6]	
UniProt	Includes over 227 million protein sequences from fully sequenced sets of viral, bacterial, archaeal and eukaryotic genomes. https://www.uniprot.org [7]	Genomics, Proteomics

Gene annotation software and databases helps reveal conserved regions of genes across organisms. Accurate prediction tools help us understand how fungal genes in genomes are structured and organized, and how evolutionary principles may affect the same sets of genes in different fungal species. These tools are certain to influence fungal biotechnology efforts, as key findings in the conservation of gene structures also assist in understanding conserved functions among fungal species [8].

With the increasing availability of fungal genomes, there has been a rapid growth in foreseeing and identifying putative Secondary Metabolites (SMs) formation genes by analyzing sequence similarities of fungal DNA compared to characterized Biosynthetic Gene Clusters (BGCs), allowing for the identification of biosynthesis genes and prediction of their functions [9,10]. Generally, accurate predictions can be made for the class of SM (e.g., polyketides, peptides, terpenes, etc.) and, in some cases, the core structures of the SMs [e.g., products of type I polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs) [11].

Botrytis cinerea-induced grey mold is widely recognized as a primary contributor to postharvest losses across a diverse spectrum of crops, encompassing fruits, vegetables, cut flowers, and flower bulbs [12]. Diseases incited by *Botrytis* species, particularly *B. cinerea*, hold a preeminent position on the global stage of plant pathology. This pre-eminence can be attributed to a combination of factors: the extraordinary breadth of host range exhibited by *Botrytis* spp., with *B. cinerea* being of prominence; the capacity to initiate quiescent infections; an impressive ability to adapt to diverse ecological niches; genetic malleability, encompassing adaptations to fungicides; the absence of host resistance mechanisms; and a versatile ecological niche transitioning from saprophyte to pathogen. One of the crucial weapons that this pathogen possess is the production of (non-specific) phytotoxic compounds to kill cells of a range of plant species [12]. In the pre-genomic era of *B. cinerea* research, eight families of SM were already isolated from in vitro mycelium. In particular, the predominant metabolites botrydial and botcinic acid were identified as two unspecific phytotoxins contributing to the necrotrophic and polyphagous lifestyle of the fungus [13].

Recognizing the pivotal role that *B. cinerea* plays among plant pathogens, its complete genome sequence has been meticulously determined [14]. The initial genome assemblies of strains B05.10 (Syngenta AG) and T4 (Genoscope) were generated using Sanger sequencing technology.

 doi:10.20944/preprints202404.0687.v1

Subsequently, Staats and van Kan improved genome assemblies of these strains using NGS Illumina technology [14]. Finally, in 2017, van Kan *et al.* reported the utilization of novel sequence data, partly derived from third-generation sequencing platforms, in combination with an optical map and a genetic map, to create a gapless, nearly complete genome assembly of the *B. cinerea* isolate B05.10 [15].

To identify the pathways involved in the production of other SMs in *B. cinerea*, a search among the genomes for genes encoding key enzymes such as NRPS (Non-Ribosomal Peptide Synthetases), PKS (PolyKetide Synthases), TS (Terpene Synthase) and DMATS (DiMethylAllylTryptophan Synthases), which are essential for the biosynthesis of peptides, polyketides, terpenes, and alkaloids, respectively [12]. Traditionally, a total of 44 genes encoding enzymes responsible for the committed biosynthetic step (called "Key" Enzyme, KE) were predicted from the sequences of the B05.10 and T4 strains [13,15].

In this work, a bioinformatic analysis was conducted using publicly available and searchable databases. The organization of candidate genes possibly related to secondary metabolism by the domains to which they belong has enabled us to compile 64 new genes possibly involved in the SM production of *B. cinerea*.

2. Results

As results derived from a bioinformatic analysis targeting the genomics of *B. cinerea*, with a specific focus on genes implicated in secondary metabolism, we utilized a synergistic approach that combined cutting-edge bioinformatic tools and genomic databases. This approach has elucidated the presence and configuration of diverse gene families associated with the biosynthesis of secondary metabolites—critical to the pathogenicity and adaptability of *B. cinerea*. This analysis not only reaffirmed the identification of previously characterized genes within the context of fungal secondary metabolism but also unveiled candidate genes hitherto undescribed, potentially involved in the synthesis of novel compounds. The results are structured and presented according to the principal gene families involved in the biosynthesis of polyketides (PKS), non-ribosomal peptides (NRPS), terpenoids, and other secondary metabolites. This organization offers a detailed overview of the complexity and richness of the genetic repertoire of *B. cinerea* dedicated to the production of these bioactive compounds.

Along manuscript we developed tables through a systematic approach to encapsulate the intricate relationship between genes of *B. cinerea* and secondary metabolism, leveraging insights from our bioinformatic analysis. As is shown in Table 3, for constructing these tables, the first column, labeled 'Name,' delineates the gene names, with newly identified genes not previously annotated or published marked by a dash to highlight their novelty. The 'ID' column follows, specifying the Gene ID as referenced within the Ensembl Fungi and FungiDB platforms, thereby providing a tangible link to the genomic underpinnings pertinent to our study. A pivotal aspect of our methodology is the color-coded 'Relationship with secondary metabolism' column, which denotes the gene's association with secondary metabolism based on database inquiries. This categorization spans from generic secondary metabolic processes to more specific involvements, such as polyketides synthases (PKS), non-ribosomal peptides (NRPS), or terpene synthases (TS), with the former two categories offering a broader context and the latter two pinpointing direct biochemical roles.

For each gene, we looked for 'Paralogous Genes', showing the results in the correspondent column, which compiles both known and yet-to-be-referenced paralogous genes, with unreferenced genes marked by their ID and known genes by their established nomenclature. This delineation underscores that the genes identified as new candidates in our tables share functional similarities with known genes, highlighting their potential roles in secondary metabolism. To ensure clarity and minimize redundancy across the gene families, we employed symbols such as asterisks (*) to indicate genes sharing identifiers with those listed undescribed and with double asterisks (**) denoting continuity without the need for repetitive listings. The notation "NA" signals the absence of paralogous genes, while additional symbols like ♣ and Δ differentiate among gene families, ensuring a coherent categorization. This meticulous table construction aims to provide a comprehensive and

 doi:10.20944/preprints202404.0687.v1

accessible overview of the genetic repertoire of *B. cinerea* related to secondary metabolism, illuminating novel genes potentially involved in its complex biosynthetic networks.

2.1. Genes Located in Domains That Code for Possible PKSs in B. cinerea

Expanding on the genomic framework of *B. cinerea* introduced earlier, this section focuses on the analysis of polyketide synthases (PKS) genes identified through bioinformatics. PKSs are essential for synthesizing a wide range of polyketides, which are crucial for the fungus's adaptability, pathogenicity, and survival [16] The diversity observed in the domains of PKS genes reflects the complexity of *B. cinerea's* secondary metabolism and its evolutionary adaptation for interacting with hosts and the environment.

Our genomic exploration revealed a substantial array of PKS genes, each harboring distinct domains critical for the polyketide production pathway. In the case of polyketide synthases, eight domains have been identified, encompassing both previously described and undescribed genes, underscoring the ongoing expansion of our understanding of PKS complexity [17]. These domains, listed in Table 2, range from the polyketide synthase dehydratase domain to the phosphopantetheine-binding domain, each playing a unique role in the enzymatic assembly and modification of polyketide chains. The domains of the dehydratase superfamily (IPR042104), acyl transferase (IPR020801), ketoreductase (IPR013968), and phosphopantetheine-binding (IPR020806) include genes that have not been described to date, in addition to those that are already known. The other domains detailed in Table 2 exclusively contain genes that have been previously identified. Notably, the *Bcchs1* gene, diverging by coding for a PKS involved in the production of pyrones, resorcilic acids, and resorcinols, showcases a thiolase-like domain (IPR016039), hinting at the presence of additional genes that could encode PKSs with similar functions.

This diversity of domains, indicative of the organism's capacity to produce a myriad of polyketide compounds, is pivotal not only for the pathogen's adaptability and survival but also for its virulence against plant hosts. The elucidation of these PKS domains and their associated genes, as detailed in Table 2, opens new avenues for targeted antifungal strategies by identifying potential vulnerabilities in the pathogen's metabolic pathways. By understanding the specific roles of different PKS domains in metabolite biosynthesis, researchers can identify potential targets for the development of novel fungicides that inhibit critical steps in polyketide synthesis, offering a blueprint for engineering disease-resistant crops.

Table 2. Domains PKS in *B. cinerea*.

Domain source	Description	Accession	Interpro code
Gene3D	Polyketide synthase, dehydratase domain superfamily	3.10.129.110	IPR042104
Pfam	Polyketide synthase, C-terminal extension	PF16197	IPR032821
Pfam	Polyketide synthase, ketoreductase domain	PF08659	IPR013968
Pfam	Polyketide synthase, dehydratase domain	PF14765	IPR020807
SMART	Polyketide synthase, acyl transferase domain	SM00827	IPR020801
SMART	Polyketide synthase, beta-ketoacyl synthase domain	SM00825	IPR020841
SMART	Polyketide synthase, dehydratase domain	SM00826	IPR020807
SMART	Polyketide synthase, phosphopantetheine-binding domain	SM00823	IPR020806

Gene3D, Pfam, and SMART are databases identifying protein domains. Gene3D categorizes structural domains. Pfam defines protein families and domains through alignments and models. SMART analyzes domain architectures and mobile domains.

 doi:10.20944/preprints202404.0687.v1

The exploration of specific PKS domains, such as those of the dehydratase superfamily, acyl transferase, ketoreductase, and phosphopantetheine-binding, has identified previously undescribed genes, suggesting uncharted pathways in polyketide biosynthesis. Expanding the known genetic repertoire of *B. cinerea* underscores the potential to discover new bioactive compounds with agricultural and pharmaceutical applications. Table 3 elaborates on the newly identified and previously unannotated genes within the PKS domains of *B. cinerea*. This table specifically focuses on detailing the genes associated with each PKS domain, shedding light on their potential roles in the biosynthesis of polyketides and their contribution to the fungus's secondary metabolic capabilities.

Table 3. Genes located in the domains that code for possible PKSs in *B. cinerea.*

Polyketide synthase, dehydratase domain superfamily. Domain:IPR042104 (Gene3D)

name	ID	Relationship with secondary metabolism				Paralogues genes	
		1	2	3	4	Undescribed	Known
Bcpks5	Bcin01g11550					Bcin03g06470; Bcin09g06360;	*Bcboa6*; *9*; *Bcpks5*; *2*; *12*; *18*; *17*; *3*; *13*; *20*; *15*; *21*;
Bcpks2	Bcin02g01680					Bcin08g02570; Bcin08g02560; Bcin12g03250*	*8*; *19*; *7*; *4*; *10*; *1*; *11*; *14*¸*16* and *Bccem1* **
Bcpks12	Bcin02g08770					NA	NA
Bcpks18	Bcin02g08830						
Bcpks17	Bcin03g02010						
Bcpks3	Bcin03g04360						
Bcpks13	Bcin03g08050						
Bcpks20	Bcin04g00640					*	**
Bcpks15	Bcin05g06220						
Bcpks21	Bcin05g08400						
Bcpks8	Bcin07g02920						
Bcpks19	Bcin08g00290						
-	Bcin09g06350					NA	NA

 doi:10.20944/preprints202404.0687.v1

Bcpks7	Bcin10g00040							
Bcpks4	Bcin11g02700					*		**
Bcpks10	Bcin13g01510							
Bcpks1	Bcin14g00600					NA		NA
Bcpks11	Bcin14g01290							
Bcpks14	Bcin16g01830							
Bcpks16	Bcin16g05040					*		**
Bcboa6	Bcin01g00060							
Bcboa9	Bcin01g00090							

Polyketide synthase, acyl transferase domain. Domain IPR020801 (SMART)

There are the genes of the last domain IPR042104, except Bcin09g06350. In addition there are the following genes as new for this domain IPR020801.

name	ID	Relationship with secondary metabolism				paralogues genes	
		1	2	3	4	Undescribed	Known
-	Bcin01g00450					NA	NA
-	Bcin04g00210			♦		Bcin01g00450	*Bcfas2*

Polyketide synthase, ketoreductase domain. Domain IPR013968 (Pfam)

There are the genes of the domain IPR042104, except *Bcpks12; 18; 17; 13; 19; 15; 14; 16;* Bcin09g06350. In addition there are the following genes as new for this domain IPR013968.

name	ID	Relationship with secondary metabolism				paralogues genes	
		1	2	3	4	Undescribed	Known

 doi:10.20944/preprints202404.0687.v1

-	Bcin03g06470			⁛		between them and Bcin08g02560; Bcin12g03250	**
-	Bcin08g02570			≠			
-	Bcin09g06360			⁛			
-	Bcin11g04550					NA	NA

Polyketide synthase, phosphopantetheine-binding domain. Domain IPR020806 (SMART)

There are the genes of the domain IPR042104, except *Bcpks18; 17; 20; 14; 16; 14; 16; Bcboa9;* Bcin09g06350 also the gene Bcin03g06470. In addition there are the following genes as new for this domain IPR020806

name	ID	Relationship with secondary metabolism				paralogues genes	
		1	2	3	4	Undescribed	Known
Bcnrps6	Bcin01g03730	♦				Bcin01g04420; Bcin01g08480; Bcin02g04610; Bcin02g06290; Bcin03g00210; Bcin03g00220: Bcin03g01550; Bcin03g01570; Bcin04g03150; Bcin06g02740; Bcin07g02750; Bcin07g02790; Bcin07g05830; Bcin08g03980; Bcin08g04860; Bcin09g02040; Bcin09g02790; Bcin10g01270; Bcin11g01420; Bcin11g02680; Bcin12g00620; Bcin12g05070; Bcin12g05180; Bcin12g05840;	*Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60* ****

 doi:10.20944/preprints202404.0687.v1

						Bcin13g02260; Bcin15g00920; Bcin15g02940; Bcin15g04320; Bcin17g00050 ***	
-	Bcin02g00016	♦		♦		Bcin01g02950; Bcin01g10250; Bcin05g07000; Bcin06g00140; Bcin06g01300; Bcin06g04410; Bcin06g04960; Bcin07g04170; Bcin08g05600; Bcin09g03540; Bcin14g03680 Bcin15g03200 ∆	*Bcfaa2* ∆∆
Bcnrps4	Bcin02g02380			♦			
-	Bcin03g00210	♦				***	****
-	Bcin03g01570						
Bclys2	Bcin04g00140	♦					
-	Bcin07g05830						
-	Bcin09g02040			♦			
-	Bcin11g01420			♦		***	****
Bcnrps8	Bcin11g02650						
Bcnrps2	Bcin12g00690	♦		♦			
Bcnrps9	Bcin14g01300	♦		♦			
Bcnrps3	Bcin16g03570	♦		♦			

Thiolase-like. Domain IPR016039 (Gene3D)

 doi:10.20944/preprints202404.0687.v1

There are the genes of the domain IPR042104, except Bcin09g06350. In addition there are the following genes as new for this domain IPR016039

name	ID	Relationship with secondary metabolism				paralogues genes	
		1	2	3	4	Undescribed	Known
Bcfas2	Bcin01g00440					Bcin01g00450 Bcin04g00210	NA
Bcpot1	Bcin01g04960			♦		Bcin03g03940; Bcin04g06330; Bcin06g05400 Bcin12g00940 *****	*Bcerg10*
-	Bcin03g03940			♦		Bcin04g06330 Bcin12g00940	*Bcpot1* and *Bcerg10* ******
-	Bcin04g04450					NA	*Bcerg13*
-	Bcin04g06330			♦		*****	******
Bcerg10	Bcin05g07430			♦		*****	*Bcpot1*
-	Bcin06g05400			♦		*****	******
-	Bcin06g07420					NA	NA
Bcerg13	Bcin11g00330					Bcin04g04450	******
-	Bcin12g00940			♦		*****	
Bccem1	Bcin16g02410					*	**

1		EnsemblFungi. Biological process: GO:0044550 secondary metabolite biosynthetic process. IEA.
	♦	Fungi DB. Metabolic pathway: Biosynthesis of various other secondary metabolites (KEGG)
2		EnsemblFungi. Biological process. GO:0009058 biosynthetic process. IEA.
3		Fungi DB. Metabolic pathway. Biosynthesis of type II polyketide backbone and products (KEGG)
	♦	Fungi DB. Metabolic pathway: aromatic polyketides biosynthesis (MetaCyc)
	⋰	NCBI: smart00823: PKS_PP; Phosphopantetheine attachment site (SMART)
	≠	Fungi DB. Uniprot. SM00822. PKS_KR (SMART)
4		Fungi DB. Metabolic pathway. Biosynthesis of siderophore group nonribosomal peptides (KEGG)

As can be observed in Table 3, following the comprehensive analysis based on the domains of these PKS genes, a total of 19 genes with domains belonging to this family have been identified, which have not yet been annotated. Specifically, 1 gene with a dehydratase domain (IPR042104), 2 genes

 doi:10.20944/preprints202404.0687.v1

with an acyl transferase domain (IPR020801), 4 genes with a ketoreductase domain (IPR013968), 6 genes with a phosphopantetheine-binding domain (IPR020806), and 6 genes with a thiolase-like domain (IPR016039) have been located. These newly identified genes allow for the expansion of the list of the 18 PKS-coding genes previously annotated for *B. cinerea* [18].

The identification of paralogous genes within the PKS domains, as meticulously cataloged in Table 3, presents an interesting glimpse into the genetic diversity and evolutionary dynamics of *B. cinerea*. For instance, the presence of multiple paralogues for the *Bcpks* genes, including *Bcpks*5 highlighted by its association with several identified paralogs such as Bcin03g06470, Bcin09g06360, and others, underpins the evolutionary strategy of gene duplication and divergence that *B. cinerea* may exploit to enhance its biosynthetic versatility and adaptability. This notion is further supported by the distinct domains these PKS genes occupy, ranging from dehydratase to phosphopantetheine-binding domains, each contributing uniquely to the polyketide synthesis pathway.

The unveiling of previously unannotated genes within the PKS domains in *Botrytis cinerea*, as presented in Table 3, enhances our comprehension of the fungus's PKS gene spectrum and hints at the existence of yet-to-be-explored metabolic routes. These routes may play crucial roles in the organism's adaptability and pathogenic prowess. The detection of such genetic variance, highlighted by paralogous gene relationships, indicates a sophisticated mechanism that allows *B. cinerea* to navigate environmental challenges effectively [19]. This adaptability, potentially leading to the emergence of novel or altered secondary metabolites, might offer the fungus distinct ecological and pathogenic advantages, including heightened virulence or increased resistance to antifungal agents [20].

This foray into the PKS domains and their paralogous gene sets not only sheds light on the secondary metabolism's complexity but also on the evolutionary strategies employed by *B. cinerea* [21]. It lays the groundwork for a deeper functional analysis of both established and novel genes, steering toward the discovery of innovative bioactive compounds. Moreover, it propels forward the development of precise antifungal interventions aimed at exploiting vulnerabilities within the pathogen's secondary metabolic pathways.

2.2. Genes Located in Domains That Code for Possible NRPSs in B. cinerea

Following the analysis of *B. cinerea's* genome, this section examines the non-ribosomal peptide synthetases (NRPS) genes identified through bioinformatics. NRPS enzymes are key in synthesizing non-ribosomal peptides, important for the fungus's adaptability and virulence. The diversity in NRPS gene domains reflects *B. cinerea's* secondary metabolic capacity and suggests an evolutionary adaptation for host and environmental interaction [22].

Our bioinformatic analysis revealed several NRPS genes, each with specific domains essential for peptide synthesis. We identified ten distinct domains, including both known and new genes, which expands our understanding of the NRPS gene family in *B. cinerea* [23]. These domains, listed in Table 4, cover functions from amino acid activation to peptide elongation and modification. For example, genes with a phosphopantetheine-binding domain play a crucial role in peptide intermediate activation and transfer.

The characterization of these NRPS domains and their genes, as presented in Table 4, contributes to a better understanding of *B. cinerea's* ability to produce diverse peptides. Understanding the roles of different NRPS domains in biosynthesis may lead to the identification of key points for intervention, offering opportunities for developing fungicides and enhancing crop resistance.

Table 4. Domains NRPS in *B. cinerea*.

Domain source	Description	Accession	Interpro code
Gene3D	ACP-like superfamily	1.10.1200.10	IPR036736
TIGRFAM	Amino acid adenylation domain	TIGR01733	IPR010071

 doi:10.20944/preprints202404.0687.v1

PROSITE patterns	AMP-binding, conserved site	PS00455	IPR020845
Pfam	AMP-dependent synthetase/ligase	PF00501	IPR000873
Gene3D	ANL, N-terminal domain	3.40.50.12780	IPR042099
Gene3D	Chloramphenicol acetyltransferase-like domain superfamily	3.30.559.10	IPR023213
Pfam	Condensation domain	PF00668	IPR001242
PROSITE patterns	Phosphopantetheine attachment site	PS00012	IPR006162
PROSITE profiles	Phosphopantetheine binding ACP domain	PS50075	IPR009081
SMART	Polyketide synthase, phosphopantetheine-binding domain	SM00823	IPR020806

Gene3D, TIGRFAM, PROSITE patterns, Pfam, and SMART are databases identifying protein domains. Gene3D categorizes structural domain families. TIGRFAM and PROSITE provide annotations for specific protein domains, focusing on functions and binding sites. Pfam defines protein families and domains through alignments and models. SMART analyzes domain architectures and genetically mobile domains.

Among the ten domains detailed in Table 4, the Condensation domain (IPR042099) stands out as the sole domain encompassing genes that have been previously characterized. The remaining domains are marked by the presence of genes that, until now, remain unannotated, unveiling a landscape ripe for exploration. This distinction underscores the potential breadth of B. cinerea's NRPS genetic repertoire, pointing to a wealth of unexplored pathways that could significantly impact the fungus's adaptability and pathogenicity. Table 5 elaborates on the genes discovered within these NRPS domains, offering insights into both recognized and novel entities. Such information enriches our comprehension of the NRPS domains, shedding light on the genetic mechanisms potentially driving the synthesis of non-ribosomal peptides crucial to Botrytis cinerea's survival and virulence.

Table 5. Genes located in the domains that code for possible NRPSs in *B. cinerea.*

ACP-like superfamily. Domain: IPR036736 (Gene 3D)

There are all the genes of the domain IPR024104 (Polyketide synthase, dehydratase domain superfamily), the domain IPR020806 (Polyketide synthase, phosphopantetheine-binding domain) and the gen Bcin09g06360 of the domain IPR013968 (Polyketide synthase, ketoreductase domain). In addition there are the following genes as new for this domain IPR036736

name	**ID**	**Relationship with secondary metabolism**				**paralogues genes**	
		1	**2**	**3**	**4**	**Undescribed**	**Known**
Bccap1	Bcin02g08010					NA	NA
-	Bcin03g01550			♦		***	****
-	Bcin06g04410					Δ	ΔΔ
-	Bcin07g01010			♦		NA	NA
-	Bcin07g02790			♦		***	****
-	Bcin15g00920			♦			
-	Bcin17g00050	♦					

 doi:10.20944/preprints202404.0687.v1

Amino acid adenylation domain. Domain IPR010071 (TIGRFAM)

There are the genes of the domain IPR020806 (Polyketide synthase, phosphopantetheine-binding domain) except Bcin03g01570; Bcin03g06470; Bcin09g02040 and Bcin11g01420. In addition there are the following genes as new for this domain IPR010071

name	ID	Relationship with secondary metabolism				paralogues genes	
		1	2	3	4	Undescribed	Known
Bcnrps5	Bcin04g01390					***	****

AMP-binding, conserved site. Domain IPR020845 (PROSITE patterns)

There are the genes *Bcnrps6, 4, 2, 9, 3; Bclys2; Bcpks5, 3, 7;* Bcin02g00016; Bcin03g00210; Bcin03g01550; Bcin03g01570; Bcin06g04410; Bcin07g01010; Bcin07g02790; Bcin07g05830; Bcin15g00920 and Bcin17g00050. In addition there are the following genes as new for this domain IPR020845

name	ID	Relationship with secondary metabolism				paralogues genes	
		1	2	3	4	Undescribed	Known
-	Bcin01g10250	♦				Δ	ΔΔ
Bcnrps7	Bcin01g11450	♦				***	****
-	Bcin02g04610						
Bcfaa2	Bcin04g00760					Δ	NA
Bcpcs60	Bcin04g01320	♦				***	****
-	Bcin04g03150			♦		NA	NA
-	Bcin05g07000					Δ	ΔΔ
-	Bcin06g00140	♦		♦			
-	Bcin06g01300						
-	Bcin06g04960	♦					
-	Bcin07g02750					***	****
-	Bcin07g04170					Δ	ΔΔ
-	Bcin08g03980	♦				***	****
-	Bcin08g04860						
-	Bcin08g05600	♦				Δ	ΔΔ
-	Bcin09g02790			♦		***	****
-	Bcin09g03540					Δ	ΔΔ
-	Bcin11g02680					***	****
Bcnrps1	Bcin12g04980	♦					
-	Bcin13g02260						
-	Bcin14g03680	♦				Δ	ΔΔ
-	Bcin15g04320	♦				***	****

 doi:10.20944/preprints202404.0687.v1

AMP-dependent synthetase/ligase. Domain IPR000873 (Pfam)

There are the genes of the domain IPR020806 (Polyketide synthase, phosphopantetheine-binding domain); except Bcin03g06470; all the genes of the domain IPR020845; *Bcpks3, 7; Bcnrps1, 5, 7*; *Bcfaa2* and *Bcpcs60.* In addition there are the following genes as new for this domain IPR000873

name	ID	Relationship with secondary metabolism				paralogues genes	
		1	2	3	4	Undescribed	Known
-	Bcin01g02950					Δ	ΔΔ
-	Bcin01g04420					***	****
-	Bcin01g08480						
-	Bcin02g06290						
-	Bcin10g00460					NA	NA
-	Bcin10g01270					***	****
-	Bcin12g00620						
-	Bcin12g05070						

ANL, N-terminal domain. Domain IPR042099 (Gene 3D)

There are all the genes of the domain IPR020806 (Polyketide synthase, phosphopantetheine-binding domain) except Bcin03g06470; the genes of the domain IPR000873 except Bcin12g05070; the genes of the domain IPR036736 except Bcin09g06360; the genes of the domain IPR020845 except *Bcfaa2*, Bcin05g07000, Bcin07g04170, Bcin08g05600. The genes *Bcpks7, 5; Bcnrps5;* Bcin03g01550; Bcin06g04410. In addition there are the following genes as new for this domain IPR042099

Name	ID	Relationship with secondary metabolism				paralogues genes	
		1	2	3	4	Undescribed	Known
-	Bcin03g00060					Bcin08g01790	NA
-	Bcin06g02740					***	****
-	Bcin08g01790					Bcin03g00060	NA

Chloramphenicol acetyltransferase-like domain superfamily. Domain IPR023213 (Gene3D)

There are the genes *Bcnrps6, 7, 4, 8, 2, 1, 9, 3; Bcpks5, 2, 3, 7* and Bcin06g04410. In addition there are the following genes as new for this domain IPR023213

Name	ID	Relationship with secondary metabolism				paralogues genes	
		1	2	3	4	Undescribed	Known

 doi:10.20944/preprints202404.0687.v1

Bcboa11	Bcin01g00110			♦		Bcin01g05970; Bcin07g07120 and Bcin15g04760 ♣	*Bcayt1*♣♣
-	Bcin01g05970			♦		♣	♣♣
-	Bcin02g02420			♦		NA	*Bccat2*
Bckgd2	Bcin02g06820			♦		Bcin11g04250	*Bcpdx1* and *Bccat2*
Bccat2	Bcin03g07910			♦		Bcin02g02420	NA
-	Bcin07g07120			♦		NA	NA
-	Bcin08g00330					NA	NA
-	Bcin11g04250	♦		♦		NA	*Bckgd2*; *Bcpdx1* and *Bcclat1*
Bclat1	Bcin12g05730			♦		Bcin11g04250	*Bckgd2* and *Bcpdx1*
Bcbot5	Bcin12g06410					NA	NA
Bcayt1	Bcin15g00050			♦		♣	*Bcboa11*
-	Bcin15g04760			♦			

Phosphopantetheine attachment site. Domain IPR006162 (PROSITE patterns)

There are the genes *Bcboa6; Bcnrps6, 4, 8, 2, 1, 9; Bcpks12, 17, 20, 15, 19, 7, 1;* Bcin03g00210, Bcin03g01550, Bcin03g06470 and Bcin07g05830.

Phosphopantetheine binding ACP domain. Domain IPR009081 (PROSITE profiles)

There are the genes *Bcboa6, 9; Bcnrps6, 4, 8,5, 2, 1, 9, 3; Bcpks5, 2, 12, 18, 17, 3, 13, 20, 15, 21, 8, 19, 4, 7, 10, 11, 1, 14, 16; Bcfas2; Bclys2;* Bcin02g00016, Bcin03g00210, Bcin03g01550, Bcin03g01570, Bcin03g06470, Bcin07g01010, Bcin07g02790, Bcin07g05830, Bcin09g02040, Bcin09g06360; Bcin11g01420; Bcin15g00920 and Bcin17g00050

Polyketide synthase, phosphopantetheine-binding domain. Domain IPR020806 (SMART)

There are the genes *Bcboa6; Bcnrps6, 4, 8, 5, 2, 9, 3; Bcpks5, 2, 12, 13, 15, 21, 8, 19, 7, 4, 10, 1, 11, 16; Bclys2;* Bcin02g00016, Bcin03g00210, Bcin03g01570, Bcin03g06470, Bcin07g05830, Bcin09g02040 and Bcin11g01420

1		EnsemblFungi. Biological process: GO:0044550 secondary metabolite biosynthetic process. IEA.
	♦	Fungi DB. Metabolic pathway: Biosynthesis of various other secondary metabolites (KEGG)
2		EnsemblFungi. Biological process. GO:0009058 biosynthetic process. IEA.
3		Fungi DB. Metabolic pathway. Biosynthesis of type II polyketide backbone and products (KEGG)
	♦	Fungi DB. Metabolic pathway: aromatic polyketides biosynthesis (MetaCyc)
4		Fungi DB. Metabolic pathway. Biosynthesis of siderophore group nonribosomal peptides (KEGG)

 doi:10.20944/preprints202404.0687.v1

As detailed in Table 5, a comprehensive analysis based on the domains of NRPS genes has identified a total of 41 genes within domains attributed to this family that have not yet been annotated. Specifically, 6 genes were located in the ACP-like superfamily domain (IPR036736), 18 genes within the AMP-binding domain (IPR020845), 8 genes in the AMP-dependent synthetase/ligase domain (IPR000873), 3 genes in the ANL, N-terminal domain (IPR042099), and 6 genes in the Chloramphenicol acetyltransferase-like domain superfamily (IPR023213). The revelation of these genes notably expands the potential repertoire of NRPS beyond the 9 previously annotated, suggesting a broader scope for secondary metabolic diversity within this pathogenic fungus [18]. This enumeration of domains harboring both recognized and newly discovered genes underscores the intricate genetic architecture that underpins the non-ribosomal peptide synthesis capability of B. cinerea, laying a foundation for future explorations into its secondary metabolism.

2.3. Genes Located in Domains That Code for Possible STCs in B. cinerea

In Section 2.3, the focus extends to genes encoding Sesquiterpene Cyclases (STCs), pivotal in the biosynthesis of sesquiterpenes, which are key contributors to B. cinerea's arsenal of bioactive compounds. These enzymes are responsible for the cyclization of farnesyl diphosphate (FPP) to yield a variety of sesquiterpenes, not only including the phytotoxic botrydial, which instigates chlorosis and necrosis in host tissues [24], but also abscisic acid, a well-documented mediator of plant stress responses and a facilitator of pathogen attack [25]. Furthermore, the recently elucidated group of 4-epi-eremophil-9-en-11-ols adds to this complex metabolic repertoire, indicating a sophisticated chemical ecology that underpins interactions with host plants [26]. The presence of these metabolites, synthesized via the STC pathways, highlights the intricate mechanisms through which B. cinerea exerts its pathogenicity and adapts to varying environmental conditions, making it a formidable adversary in agricultural settings [18].

Table 6 delineates the domains associated with STCs in B. cinerea, emphasizing the diversity within this enzyme family. Notably, a singular domain, IPR008949, harbors previously characterized STCs alongside genes yet to be annotated. This domain's inclusion of 13 genes, with specific subsets sharing the IPR000092 and IPR024652 domains, underscores the complexity and potential expansiveness of the STC repertoire in B. cinerea. The detailed enumeration of these domains invites further investigation into their roles and contributions to the fungal secondary metabolism.

Table 6. Domains STC in *B. cinerea*.

Domain source	Description	Accession	Interpro code
Gene3D	Isoprenoid synthase domain superfamily		IPR008949
Pfam	Polyprenyl synthetase		IPR000092
Pfam	Trichodiene synthase		IPR024652

Gene3D and Pfam are databases identifying protein domains. Gene3D categorizes structural domain families. Pfam defines protein families and domains through alignments and models.

As detailed in Table 7, a comprehensive domain-based analysis has led to the identification of three genes associated with STC domains. Specifically, three genes have been identified within the Isoprenoid synthase domain superfamily (IPR008949), one gene within the Polyprenyl synthetase domain (IPR000092), and one gene within the Trichodiene synthase domain (IPR024652). Notably, the latter two genes are also classified under the domain IPR008949, bringing the total to three new genes identified as part of this family. This expansion of the potential gene pool represents a significant extension beyond the currently characterized STCs in *B. cinerea*, paving the way for further exploration into previously unexplored facets of its secondary metabolism.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 April 2024 doi:10.20944/preprints202404.0687.v1

Table 7. Genes located in the domains that code for possible STCs in B. cinerea.

Isoprenoid synthase domain superfamily. Domain: IPR008949 (Gene 3D)							
Name	**ID**	**Relationship with secondary metabolism**				**paralogues genes**	
		1	**2**	**3**	**4**	**Undescribed**	**Known**
Bcstc5	Bcin01g03520					NA	NA
Bcphs1	Bcin01g04560						
Bccoq1	Bcin02g05540		♦			Bcin14g01170	NA
Bcstc4	Bcin04g03550						*Bcstc3*
Bcpax1	Bcin05g05670		♦				
Bcerg9	Bcin06g02400						
Bcstc2	Bcin08g02350						
-	Bcin08g03510						NA
-	Bcin11g06510					NA	
Bcbot2 /stc1	Bcin12g06390						
Bcstc3	Bcin13g05830						*Bcstc4*
-	Bcin14g01170		♦				*Bccoq1*
Bcerg20	Bcin15g03120		♦				NA

Polyprenyl synthetase. Domain IPR000092 (Pfam)
There are the genes *Bccoq1*, *Bcpax1*, *Bcerg20* and Bcin14g01170

Trichodiene synthase. Domain IPR024652 (Pfam)
There are the genes *Bcstc2* and Bcin11g06510

1		EnsemblFungi. Biological process: GO:1901362 organic cyclic compound biosynthetic process. IEA.
2		EnsemblFungi. Biological process. GO:0009058 biosynthetic process. IEA.
	♦	EnsemblFungi. Biological process. GO:0008299 isoprenoid biosynthetic process
3		NCBI: cl00210 Isoprenoid Biosynthesis enzymes, Class 1.
4		Fungi DB. Uniprot. SSF48576. Terpenoid synthases superfamily (SUPFAM)

2.4. Genes Located in Domains That Code for Possible DTCs in B. cinerea

In Section 2.4, attention is turned towards the characterization of genes associated with Diterpene Cyclases (DTCs), a class of enzymes integral to the biosynthesis of diterpenes. Diterpenes are a diverse group of secondary metabolites known for their role in plant-microbe interactions, including defence mechanisms against phytopathogens. In *B. cinerea*, the exploration of DTCs

 doi:10.20944/preprints202404.0687.v1

illuminates new aspects of its secondary metabolism, potentially revealing pathways involved in the synthesis of diterpenoid compounds [27]. The intricate nature of these pathways underscores the fungus's sophisticated adaptability and pathogenic arsenal.

In the comprehensive analysis of *B. cinerea's* diterpene cyclases (DTCs) within the IPR008930 domain, a detailed examination has unveiled five genes, with *Bcdtc1* being particularly notable for its well-established diterpene cyclase activity. This discovery, as presented in Table 8, emphasizes *Bcdtc1*'s distinct role within the domain. It's important to note, however, that while *Bcdtc2* and *Bcdtc3* have traditionally been recognized as part of the secondary metabolism gene set in *B. cinerea*, current investigations reveal that Bcdtc2 no longer has available information, and Bcdtc3 does not align with any domain typically associated with this family [18]. Despite this, the domain IPR008930 encompasses five genes, inclusive of Bcdtc1, underscoring a specialized contribution towards diterpene synthesis in *B. cinerea*. Furthermore, this analysis has led to the discovery of an additional, yet unannotated, gene (Bcin02g00670), emerging as a new candidate within the scope of the fungus's secondary metabolism.

Table 8. Genes located in the domains that code for possible DTCs in *B. cinerea*.

Terpenoid cyclases/protein prenyltransferase alpha-alpha toroid. Domain: IPR008930 (Superfamily)							
name	**ID**	**Relationship with secondary metabolism**				**paralogues genes**	
		1	**2**	**3**	**4**	**Undescribed**	**Known**
Bccdc43	Bcin01g04020						*Bcram1* and *Bcbet2*
Bcdtc1	Bcin01g04920						*Bcdtc3*
-	Bcin02g00670	Ξ	#			NA	NA
Bcram1	Bcin03g06350						*Bccdc43* and *Bcbet2*
Bcbet2	Bcin06g05320			♦			NA

1	Ξ	EnsemblFungi. Biological process: GO:0006694 steroid biosynthetic process. IEA.
2	#	EnsemblFungi. Biological process. GO:0016104 triterpenoid biosynthetic process. IEA.
3		NCBI: cd02895: Geranylgeranyltransferase types I
	♦	NCBI: cd02894: Geranylgeranyltransferase type II
		NCBI: cl27572: Terpene synthase, N-terminal domain
		NCBI: cd02892: Squalene cyclase (SQCY) domain subgroup 1
		NCBI: cd02893: Protein farnesyltransferase (FTase)
4		Fungi DB. Uniprot. SSF48239 Terpenoid cyclases/Protein prenyltransferases (SUPFAM)

2.5. Genes Located in Domains That Code for Possible DMATSs in B. cinerea

In our pursuit of genes coding for the two known DiMethylAllylTryptophan Synthases in *B. cinerea*, we have discerned that these genes coalesce within two distinct domains, IPR017795 and IPR033964. Notably, within the latter, a hitherto unidentified gene emerges, augmenting our insights into the genetic landscape of this phytopathogen. The encapsulation of this discovery is delineated in Table 9, signifying a pivotal expansion of our knowledge concerning the enzymatic repertoire of this fungus.

Table 9 showcases the DMATS domains identified in *B. cinerea*, highlighting two main domains: IPR017795 (Aromatic prenyltransferase DMATS-type) and IPR033964 (Aromatic prenyltransferase).

 doi:10.20944/preprints202404.0687.v1

It lists known genes, *Bcdmats*2 and *Bcdmats*1, alongside a new, unannotated gene found within the IPR033964 domain. This table serves as a critical update to the DMATS gene inventory for the B05.10 strain of this fungus.

Table 9. Genes located in the domains that code for possible DMATS in *B. cinerea*.

Aromatic prenyltransferase, DMATS-type. Domain: IPR017795 (Pfam)

Name	ID	Relationship with secondary metabolism			paralogues genes	
		1	2	3	Undescribed	Known
Bcdmats2	Bcin14g04900				NA	NA
Bcdmats1	Bcin16g01940					

Aromatic prenyltransferase. Domain: IPR033964 (SFLD)

There are the genes *Bcdmats2* and *Bcdmats1*

Name	ID	Relationship with secondary metabolism			paralogues genes	
		1	2	3	Undescribed	Known
-	Bcin06g02600	♦	♦		NA	NA

1		EnsemblFungi. Biological process: GO:0009820 alkaloid metabolic process. IEA.
	♦	EnsemblFungi. Molecular function. GO:0004659 prenyltransferase activity. IEA.
2		NCBI: cd13929: aromatic prenyltransferases (PTases) of the DMATS/CymD familiy
	♦	NCBI: pfam11468: Aromatic prenyltransferase Orf2
3		Fungi DB. Uniprot. PTHR40627. Indole prenyltransferase tdib-related (PANTHER)

The scrutiny of DMATS domains as articulated in Table 9 has unveiled the presence of an additional, previously unannotated gene within the IPR033964 domain. This discovery not only enriches our list of DMATS genes in *B. cinerea* but also alludes to the potential existence of untapped metabolic pathways that could play a pivotal role in the organism's adaptability and pathogenic capabilities [18].

2.6. New Candidate Genes Possibly Related to the Secondary Metabolism of B. cinerea

In consolidating our findings from the comprehensive bioinformatics analyses of B. cinerea's genome, we have unveiled an expanded repertoire of genes implicated in the synthesis of secondary metabolites. Through this study's journey across the diverse enzymatic landscapes of PKS, NRPS, STC, DTC, and DMATS, we have illuminated the genetic underpinnings contributing to the pathogen's complexity and virulence. The compilation of newly identified genes in Table 10 represents a pivotal enhancement of the known gene pool associated with these critical families. Moreover, the classification into groups such as NRPS, PKS-NRPS hybrids, and potentially other

 doi:10.20944/preprints202404.0687.v1

hybrid forms, as informed by the current databases, introduces a layer of complexity in precisely defining these genes' roles. This nuance highlights the evolving nature of our understanding, acknowledging that some genes might exhibit functionalities of one or more enzyme types related to secondary metabolism (SM). The elucidation of these genes not only broadens our understanding of B. cinerea's metabolic capacities but also underscores the dynamic evolution of its genome in adapting to various ecological niches and host interactions.

Table 10. New candidate genes possibly related to the secondary metabolism of *B. cinerea*.

PKS	
ID	**Observation**
Bcin01g00450	Domain IPR020801. Paralogue gene *Bcfas2*. Metabolic pathway: Biosynthesis of type II polyketide backbone and products.
Bcin03g03940	Domain IPR016039. Paralogues genes *Bcpot1* and *Bcerg10*. Metabolic pathway: aromatic polyketides biosynthesis.
Bcin03g06470	Domains IPR013968 and IPR020806. Paralogues genes *Bcboa6*; *9*; *Bcpks5*; *2; 12*; *18*; *17*; *3; 13; 20*; *15*; *21*; *8*; *19; 7*; *4*; *10; 1; 11; 14¸16* and *Bccem1*.
Bcin04g00210	Domain IPR020801. Paralogue gene *Bcfas2*. Metabolic pathway: aromatic polyketides biosynthesis. Bikaverin biosynthesis (MetaCyc).
Bcin04g04450	Domain IPR016039. Paralogue gene *Bcerg13*.
Bcin04g06330	Domain IPR016039. Paralogues genes *Bcpot1* and *Bcerg10*. Metabolic pathway: aromatic polyketides biosynthesis.
Bcin06g05400	
Bcin06g07420	Domain IPR016039.
Bcin08g02570	Domain IPR013968. Paralogues genes *Bcboa6*; *9*; *Bcpks5*; *2; 12*; *18*; *17*; *3; 13; 20*; *15*; *21*; *8*; *19; 7*; *4*; *10; 1; 11; 14¸16* and *Bccem1*.
Bcin09g06350	Domain IPR042104.
Bcin09g06360	Domains IPR013968 and IPR036736. Paralogues genes *Bcboa6*; *9*; *Bcpks5*; *2; 12*; *18*; *17*; *3; 13; 20*; *15*; *21*; *8*; *19; 7*; *4*; *10; 1; 11; 14¸16* and *Bccem1*.
Bcin11g04550	Domain IPR013968.
Bcin12g00940	Domain IPR016039. Paralogues genes *Bcpot1* and *Bcerg10*. Metabolic pathway: aromatic polyketides biosynthesis.
PKS-NRPS hybrids	
Bcin02g00016	Domain IPR020806. Paralogue gene *Bcfaa2*. Metabolic pathway: aromatic polyketides biosynthesis and biosynthesis of siderophore group nonribosomal peptides.
Bcin07g01010	Domain IPR036736. Metabolic pathway: aromatic polyketides biosynthesis and biosynthesis of siderophore group nonribosomal peptides.
NRPS or PKS-NRPS hybrids	

 doi:10.20944/preprints202404.0687.v1

Bcin01g10250	Domain IPR020845. Paralogue gene *Bcfaa2*. Metabolic pathway: Biosynthesis of type II polyketide backbone and products.
Bcin01g02950	Domain IPR000873. Paralogue gene *Bcfaa2*.
Bcin02g04610	Domain IPR020845. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60*.
Bcin05g07000	Domain IPR020845. Paralogue gene *Bcfaa2*.
Bcin06g01300	
Bcin07g02750	Domain IPR020845. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60*.
Bcin07g04170	Domain IPR020845. Paralogue gene *Bcfaa2*.
Bcin08g00330	Domain IPR023213.
Bcin08g01790	Domain IPR042099.
Bcin08g04860	Domain IPR020845. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60*.
Bcin09g03540	Domain IPR020845. Paralogue gene *Bcfaa2*.
Bcin11g02680	Domain IPR020845. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60*.
Bcin13g02260	

NRPS or PKS or PKS-NRPS hybrids

Bcin01g05970	Domain IPR023213. Paralogue gene *Bcayt1*. Metabolic pathway: aromatic polyketides biosynthesis.
Bcin01g04420	Domain IPR000873. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60*.
Bcin01g08480	
Bcin02g02420	Domain IPR023213. Paralogue gene *Bccat2*. Metabolic pathway: aromatic polyketides biosynthesis.
Bcin02g06290	Domain IPR000873. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60*.
Bcin03g00060	Domain IPR042099.
Bcin03g01550	Domain IPR036736. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60*. Metabolic pathway: aromatic polyketides biosynthesis.
Bcin03g01570	Domain IPR020806. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60*.
Bcin03g00210	Domain IPR020806. Paralogue gene *Bcfaa2*.
Bcin04g03150	Domain IPR020845. Metabolic pathway: aromatic polyketides biosynthesis.
Bcin06g00140	Domain IPR020845. Paralogue gene *Bcfaa2*. Metabolic pathway: aromatic polyketides biosynthesis.
Bcin06g02740	Domain IPR042099. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60*.
Bcin06g04410	Domain IPR036736. Paralogue gene *Bcfaa2*.
Bcin06g04960	Domain IPR020845. Paralogue gene *Bcfaa2*. Metabolic pathway: Biosynthesis of type II polyketide backbone and products.
Bcin07g02790	Domain IPR036736. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60*. Metabolic pathway: aromatic polyketides biosynthesis.
Bcin07g05830	Domain IPR020806. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60*.
Bcin07g07120	Domain IPR023213. Metabolic pathway: aromatic polyketides biosynthesis.

 doi:10.20944/preprints202404.0687.v1

Bcin08g03980	Domain IPR020845. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60*. Metabolic pathway: Biosynthesis of type II polyketide backbone and products.
Bcin08g05600	Domain IPR020845. Paralogue gene *Bcfaa2*. Metabolic pathway: Biosynthesis of type II polyketide backbone and products.
Bcin09g02040	Domain IPR020806. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60*. Metabolic pathway: aromatic polyketides biosynthesis.
Bcin09g02790	Domain IPR020845. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60*. Metabolic pathway: aromatic polyketides biosynthesis.
Bcin10g00460	Domain IPR000873. Metabolic pathway: Biosynthesis of type II polyketide backbone and products.
Bcin10g01270	Domain IPR000873. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60.*
Bcin11g01420	Domain IPR020806. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60*. Metabolic pathway: aromatic polyketides biosynthesis.
Bcin11g04250	Domain IPR023213. Paralogues genes *Bckgd2; Bcpdx1* and *Bcclat1*. Metabolic pathway: aromatic polyketides biosynthesis.
Bcin12g00620	Domain IPR000873. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60.*
Bcin12g05070	
Bcin14g03680	Domain IPR020845. Paralogue gene *Bcfaa2*. Metabolic pathway: Biosynthesis of type II polyketide backbone and products.
Bcin15g00920	Domain IPR036736. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60.* Metabolic pathway: aromatic polyketides biosynthesis.
Bcin15g04320	Domain IPR020845. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60*. Metabolic pathway: Biosynthesis of type II polyketide backbone and products.
Bcin15g04760	Domain IPR023213. Paralogue gene *Bcboa11*. Metabolic pathway: aromatic polyketides biosynthesis.
Bcin17g00050	Domain IPR036736. Paralogues genes *Bcnrps7; 4; 5; 8; 2; 1; 9; 3; Bclys2* and *Bcpcs60.*

STC

Bcin08g03510	Domain IPR008949. Isoprenoid Biosynthesis enzymes, Class 1. Terpenoid synthases superfamily.
Bcin11g06510	Domains IPR008949 and IPR024652. Isoprenoid Biosynthesis enzymes, Class 1. Terpenoid synthases superfamily. Identified as ***Bstc7*.**
Bcin14g01170	Domains IPR008949 and IPR000092. Paralogue gene *Bccoq1*. Biological process: isoprenoid biosynthetic process. Terpenoid synthases superfamily.

DTC

 doi:10.20944/preprints202404.0687.v1

Bcin02g00670	Domain IPR008930. Squalene cyclase (SQCY) domain subgroup 1. Terpenoid cyclases/Protein prenyltransferases.

DMATS

Bcin06g02600	Domain IPR033964. Prenyltransferase activity. IEA. Aromatic prenyltransferase Orf2.

In this updated examination of B. cinerea's secondary metabolism genes, critical adjustments and noteworthy additions have been made to the gene catalog. Particularly, the Bcdtc2 gene has been excluded from our considerations due to its absence in recent database records. Additionally, the Bcbik genes, once presumed relevant to secondary metabolism, are identified as non-functional for the B05.10 strain, thereby not contributing to bikaverin production [28]. Furthermore, the Bcphs1 gene designation revealed two genes, Bcin08g03790 and Bcin01g04560, with the former being a novel discovery warranting further investigation due to its potential involvement in retinal production, unlike the previously studied Bcin01g04560 [29].

Most notably, the gene Bcin11g06510, now classified as Bcstc7, represents a significant addition to the B. cinerea gene repertoire. Initial transcriptomic, metabolomic, and phenotypic characterizations suggest Bcstc7 as a key enzyme in the biosynthesis of eremophilenol, marking the first documentation of this biosynthetic pathway in the fungus [26]. The comprehensive list in Table 10, marking the culmination of our search, not only enhances the catalog of known secondary metabolism genes but also sets the stage for future research endeavors aimed at unraveling the metabolic complexity and ecological strategies of this phytopathogen.

3. Discussion

In the comprehensive investigation presented within this study, we embarked on a detailed exploration of the secondary metabolism of *B. cinerea*, focusing specifically on the genetic underpinnings of its biosynthetic capabilities. Through the meticulous bioinformatic analysis of known and novel gene families—namely Polyketide Synthases (PKS), Non-Ribosomal Peptide Synthetases (NRPS), Sesquiterpene Cyclases (STC), Diterpene Cyclases (DTC), and DiMethylAllylTryptophan Synthases (DMATS)—this research has significantly advanced our understanding of the molecular basis for the diverse array of secondary metabolites produced by this phytopathogenic fungus. The localization of these genes provides not only a deeper insight into the metabolic complexity of *B. cinerea* but also illuminates the evolutionary strategies it employs to thrive within its ecological niche and against its plant hosts.

The findings reported herein contribute to a growing body of evidence that underscores the dynamic nature of fungal secondary metabolism, revealing an intricate network of genetic elements that drive the synthesis of compounds essential for pathogenicity, survival, and competition. The identification of a multitude of previously unannotated genes across these key enzymatic families opens new avenues for research into the specific roles these genes play in the life cycle of *B. cinerea*, their potential impact on host-pathogen interactions, and their implications for agricultural management practices.

The elucidation of new Polyketide Synthase (PKS) genes within the genome of *B. cinerea* marks a significant advancement in our understanding of the fungal arsenal against plant hosts. The identification of these genes not only enriches the catalog of known PKS in this phytopathogen but also provides insights into the potential for an expanded repertoire of polyketides, compounds known for their critical roles in fungal virulence and adaptability. The presence of these newly discovered PKS genes suggests an evolutionary advantage for *B. cinerea*, enabling the fungus to synthesize a broader array of polyketides that may contribute to its pathogenicity and survival in diverse environmental conditions [17].

 doi:10.20944/preprints202404.0687.v1

Comparatively, this expansion mirrors findings in other phytopathogenic fungi, where the diversity of PKS genes correlates with the organism's ability to infect its host and evade plant defenses. For instance, studies in fungi such as *Fusarium graminearum* and *Aspergillus nidulans* have demonstrated the pivotal role of PKS-derived metabolites in host-pathogen interactions, underscoring the evolutionary pressure on these organisms to diversify their metabolic capabilities. Our findings in *B. cinerea* add to this narrative, suggesting that the expansion of the PKS gene family is a conserved strategy among fungi to enhance their pathogenic toolkit [30]. Moreover, the discovery of these PKS genes in *B. cinerea* not only has implications for our understanding of fungal pathogenesis but also highlights potential targets for the development of novel antifungal strategies. By elucidating the functions and products of these newly identified genes, future research may pave the way for innovative approaches to control *B. cinerea* infections, potentially reducing the agricultural and economic impact of this pathogen [19].

The advancement in genomic exploration has led to the significant identification of an expanded repertoire of Non-Ribosomal Peptide Synthetase (NRPS) genes in *B. cinerea*. This expansion not only augments our comprehension of the NRPS gene family within this formidable phytopathogen but also sets the stage for a deeper investigation into the myriad of secondary metabolites these genes are poised to produce. The NRPS enzymes are quintessential for the synthesis of a wide range of non-ribosomal peptides, many of which play pivotal roles in the pathogenicity and environmental adaptability of fungi. The diversity of NRPS genes uncovered in *B. cinerea* hints at a sophisticated arsenal capable of synthesizing diverse bioactive compounds, potentially contributing to the fungus's success as a pathogen across a broad spectrum of host plants [23]. The functional diversity suggested by the newly identified NRPS domains in *B. cinerea* speaks to the evolutionary pressure on this pathogen to innovate its metabolic pathways. These peptides, often characterized by their complex structures and specific biological activities, can include toxins, siderophores, and immunosuppressants, among others, each contributing uniquely to the pathogen's interaction with its host and survival within competitive microbial ecologies. For example, the production of siderophores, which sequester iron from the host or environment, showcases a direct link between NRPS activity and fungal virulence, underscoring the critical role these enzymes play in pathogen survival and proliferation [31].

The identification of novel NRPS genes aligns with findings in other pathogenic fungi, where the variability and abundance of these genes have been correlated with a heightened capacity for host invasion and colonization. This suggests that the NRPS gene family's expansion within *B. cinerea* may be reflective of similar evolutionary adaptations, enabling a more versatile and robust engagement with host defenses. Such an adaptive mechanism is indicative of the pathogen's evolved strategies to ensure survival and pathogenic success, hinting at a complex interplay between fungal metabolism and host defense mechanisms. Additionally, the discovery of these genes opens new avenues for antifungal drug discovery. By targeting the unique enzymatic mechanisms of NRPSs, novel therapeutic strategies can be developed to mitigate the impact of *B. cinerea* on agriculture without relying on traditional fungicides that pose risks of resistance development and environmental harm. This strategic approach towards antifungal intervention underscores the critical importance of understanding the functional diversity and biological roles of NRPSs in *B. cinerea's* lifecycle and pathogenicity.

The integration of Sesquiterpene Cyclases (STC) in our study provides critical insights into the biosynthesis of sesquiterpenes in *B. cinerea* and their integral role in the interaction between pathogen and host. Sesquiterpenes, known for their diverse and complex structures, play significant roles in fungal defense mechanisms and in mediating interactions with plant hosts. The identification of new STC genes in *B. cinerea* not only enhances our understanding of the sesquiterpene biosynthetic pathways but also offers clues to the chemical diversity employed by this fungus to adapt and thrive in various ecological niches. This expanded genetic repertoire suggests a sophisticated mechanism by which *B. cinerea* can produce sesquiterpenes that contribute to its virulence and ability to overcome plant defense systems, indicating a critical area for future research on pathogen-host dynamics [26,27]. Similarly, the exploration of Diterpene Cyclases (DTC) genes within *B. cinerea* uncovers new

 doi:10.20944/preprints202404.0687.v1

dimensions of the fungus's ability to synthesize diterpenes, compounds less characterized in fungal secondary metabolism yet known for their potential roles in fungal development and host interactions. The discovery of DTC genes paves the way for understanding the biosynthesis of diterpenes in *B. cinerea*, shedding light on their functions and implications in fungal physiology and pathogenicity. Intriguingly, the diterpenes' scarce representation in the known metabolome of *B. cinerea* raises questions about their roles and prevalence within this fungal species. This gap in knowledge opens up avenues for intensive research, suggesting that a deeper investigation into the DTC-mediated diterpene biosynthesis could reveal novel metabolites with significant ecological and pathogenic relevance. The potential of these diterpenes to act as virulence factors or modulators of plant immune responses makes them compelling subjects for future studies, aiming to elucidate their contributions to the sophisticated interplay between *B. cinerea* and its host plants [25].

The recent discovery of a new DiMethylAllylTryptophan Synthase (DMATS) gene within the *B. cinerea* genome represents a significant milestone in fungal secondary metabolism research. DMATS enzymes are key players in the biosynthesis of indole-derived secondary metabolites, which are pivotal for various biological processes including microbial competition, plant-microbe interactions, and pathogenesis. The identification of this gene provides new outlooks for understanding the biosynthetic capabilities of *B. cinerea*, offering insights into previously unexplored routes leading to the production of novel indole derivatives. These compounds could have profound implications for the pathogen's adaptability, virulence, and resistance mechanisms against plant defense compounds [13].

The potential biosynthetic pathways and products mediated by this newly identified DMATS gene could significantly enhance our comprehension of the chemical ecology of *B. cinerea*. Given the crucial roles that indole-derived metabolites play in fungal pathogenicity and host interaction, elucidating these pathways may reveal novel targets for antifungal intervention. Indeed, the modulation of such pathways offers a promising avenue for the development of innovative crop protection strategies. By targeting specific enzymes or steps within these pathways, it might be possible to mitigate the virulence of *B. cinerea* without affecting beneficial microbes, thereby reducing the reliance on broad-spectrum fungicides that pose risks of resistance development and environmental damage [32].

Furthermore, the exploration of DMATS-mediated biosynthetic routes underscores the potential for discovering new natural products with applications beyond agriculture, including pharmaceuticals and industrial biotechnology. The versatility of indole derivatives as bioactive compounds highlights the importance of this gene discovery not only for plant pathology and crop protection but also for harnessing the synthetic potential of fungal secondary metabolites [33].

Throughout this comprehensive study, we have significantly expanded our understanding of the secondary metabolism of *B. cinerea* by identifying and characterizing key genes within the PKS, NRPS, STC, DTC, and DMATS gene families. These findings not only broaden the genetic repertoire associated with the production of secondary metabolites in this fungal pathogen but also underscore the complexity and diversity of mechanisms *B. cinerea* utilizes to interact with its hosts and adapt to varying environments. The elucidation of these genes lays a solid foundation for future research aimed at understanding the specific biosynthetic pathways and metabolic products involved in fungal pathogenicity and resistance to plant defenses [13,19] .

From a biotechnological and crop disease management perspective, these discoveries hold vast potential for the development of novel and more effective disease control strategies. Manipulating these metabolic pathways through advanced genetic tools could allow to produce specific bioactive compounds with applications in agriculture, as well as in the pharmaceutical and chemical industries. Moreover, detailed knowledge of *B. cinerea's* virulence mechanisms pave the way for the creation of plant cultivars with enhanced resistance, thus reducing dependency on chemical fungicides and contributing to more sustainable and environmentally friendly agricultural systems [34]. Looking forward, it is evident that there is still much to uncover about the complex secondary metabolism of *B. cinerea*. Future research prospects in this field include detailed exploration of the newly identified biosynthetic pathways, functional characterization of unknown metabolic products,

 doi:10.20944/preprints202404.0687.v1

and investigation of the specific molecular interactions between the pathogen and its hosts. Additionally, advancing our understanding of the genetic and epigenetic regulation of these biosynthetic systems to develop targeted and tailored interventions is essential. Continuing to expand our knowledge in these areas will not only enrich fundamental science in mycology and plant pathology but also have significant practical applications in crop protection and biotechnology.

4. Materials and Methods

4.1. Genomic Database Capabilities

For our research, we utilized the extensive genomic repositories of Ensembl Fungi and FungiDB to investigate the genetic framework of *Botrytis cinerea*. Ensembl Fungi is a crucial genomic browser that consolidates data from major nucleotide sequence databases, enabling its visualization through its own genomic browser and providing access via Perl and RESTful APIs, MySQL databases, and FTP sites for data in various formats such as FASTA, EMBL, and GTF [35]. Currently, it hosts 1,505 genomes, offering updated annotations on genome alignments, interactions among fungal genes, and protein features for all included species. Ensembl Fungi stands out for its collaborative approach, integrating data from multiple international sources and offering advanced tools for the analysis and visualization of genomic sequences and expression data [36,37]. This resource is invaluable for identifying genes involved in secondary metabolism, providing a robust platform for the genetic exploration of fungal pathogens and their interaction with the environment [38]. The application of genomic data to study mycotoxin synthesis and function, pathogenesis, and other aspects of fungal biology is still in its infancy, with preliminary examination of microarray data collected from *Fusarium verticillioides* liquid cultures providing evidence of widespread differential gene expression over time [39] (Brown et al., 2006).

FungiDB, a core component of the EuPathDB platform, plays a pivotal role in the realm of functional genomics research for fungi and oomycete species [40]. This comprehensive database integrates a wide array of data types, including genomic, transcriptomic, proteomic, and phenotypic information, from a diverse spectrum of organisms. With a repository that encompasses nearly 100 genomes from multifarious origins, FungiDB stands as a premier resource for researchers. It offers an intuitive interface equipped with sophisticated bioinformatics tools designed to support custom, in-depth in silico analyses [41]. These tools enable researchers to perform tailored experiments and analyses directly within the platform, facilitating a seamless research workflow.

Beyond its user-friendly design, FungiDB provides an advanced Galaxy-based workspace that empowers users to construct and execute custom analysis pipelines [42]. This feature is particularly beneficial for conducting comprehensive data analyses, such as RNA sequencing (RNA-Seq) and variant calling, allowing for the nuanced exploration of genetic sequences and expression patterns. Moreover, FungiDB's integration with EuPathDB highlights its commitment to supporting the fungal research community by providing access to a wealth of genomic resources. This collaboration enhances the database's utility, making it an indispensable tool for researchers aiming to conduct comparative genomics studies, understand evolutionary relationships, and develop targeted interventions for fungal diseases [43]. In this way, FungiDB complements the genomic data and analytical tools provided by Ensembl Fungi, together offering a comprehensive suite of resources for the genetic exploration of fungal pathogens and their interactions with hosts and environments.

FungiDB enhances our research capabilities by offering a rich, integrated platform for accessing and analyzing fungal genomic data [40,41]. Its contribution to the field of functional genomics is invaluable, providing researchers with the tools needed to delve into the genetic intricacies of *Botrytis cinerea* and other fungal species, enabling the identification of key genetic markers and pathways involved in pathogenesis, resistance, and secondary metabolism [44].

4.2. Bioinformatic Analysis of Secondary Metabolism Genes

To delve into the secondary metabolism of *Botrytis cinerea* (B05.10), we employed a methodical bioinformatic strategy leveraging the Ensembl Fungi and FungiDB databases, supplemented by

 doi:10.20944/preprints202404.0687.v1

NCBI resources. Our initial step involved the meticulous identification of genes previously associated with *B. cinerea*'s secondary metabolism, cataloged by their specific Gene IDs. This approach is supported by the comprehensive platform for gene expression analysis in *B. cinerea* described by Aguayo & Canessa (2022) [45], which has been instrumental in identifying and analysing genes involved in secondary metabolism.

a) Identifying Genomic Domains and Features: Utilizing the Ensembl Fungi database, we systematically explored the "Domains & Features" section for each identified gene. This exploration unveiled the distinct domains associated with each gene, facilitating a comprehensive understanding of their potential functions (Supplementary Material - Figure S1). Subsequently, we navigated through each domain to uncover additional *B. cinerea* genes classified within these domains, thereby expanding our candidate gene pool (Supplementary Material - Figure S2). This step aligns with methodologies outlined in global and proteome-wide analyses by Zhang et al. (2020) [46] and Xu et al. (2020) [47], demonstrating the relevance of bioinformatic strategies in elucidating complex metabolic pathways.
b) Database Cross-Referencing and Functional Analysis: Beyond Ensembl Fungi, our search extended to FungiDB and NCBI. This cross-referencing enabled us to consult various parameters related to secondary metabolism, including predicted functions and metabolic pathways. Specifically, we scrutinized biological processes, molecular functions, and potential paralogous relationships, aiming to broaden our collection of candidate genes involved in secondary metabolism (Supplementary Material - Figure S3). The identification and analysis of miRNAs and siRNAs in *B. cinerea* by Liu et al., (2022) [47] further support our cross-referencing and functional analysis steps.
c) Compilation and Categorization: The culled information was methodically organized into tables, categorizing genes into pertinent groups based on their association with secondary metabolism pathways such as polyketide synthases (PKS), non-ribosomal peptide synthetases (NRPS), and others. Each table (e.g., Table 3 for PKS, Table 5 for NRPS) presented a consolidated view of genes implicated in distinct biosynthetic mechanisms (Tables 3, 5-8).
d) Analytical Approach: Our bioinformatic analysis was characterized by a granular examination of each gene. We juxtaposed the biological processes, molecular functions, and paralogous gene relationships against the backdrop of their domain affiliations. This thorough analysis, enriched by additional insights from the databases, allowed us to confidently assign genes to specific biosynthetic categories (PKS, NRPS, etc.), laying the groundwork for proposing new candidate genes potentially involved in *B. cinerea's* secondary metabolism (Table 9).

This bioinformatic framework not only facilitated the identification of new genes potentially implicated in the secondary metabolism of *B. cinerea* but also underscored the complexity and diversity of fungal biosynthetic pathways.

5. Conclusions

This study represents a significant advancement in our understanding of the secondary metabolism of *B. cinerea*, uncovering an expanded spectrum of candidate genes implicated in the synthesis of secondary metabolites. The thorough integration of bioinformatics analyses with available genomic information has shed light on the genetic complexity underlying the pathogenicity and adaptability of this fungus. By exploring the enzymatic families of PKS, NRPS, STC, DTC, and DMATS, we have identified previously unannotated genes, offering new insights into the diversity and evolution of *B. cinerea's* genome.

The identification of these genes not only enriches our knowledge of *B. cinerea's* metabolic capabilities but also highlights the dynamic evolution of its genome, facilitating adaptations to diverse ecological niches and host interactions. The discovery of PKS-NRPS hybrid genes further underscores the complexity of this fungus's secondary metabolism, paving the way for the exploration of yet-unknown metabolic pathways. These findings are pivotal for advancing our understanding of the mechanisms through which *B. cinerea* interacts with its hosts and survives in varied environments.

 doi:10.20944/preprints202404.0687.v1

The expansion of the genetic repertoire associated with the synthesis of secondary metabolites in *B. cinerea* carries significant implications for agricultural management and biotechnology. Manipulating these metabolic pathways, using advanced genetic tools, could lead to the production of specific bioactive compounds with applications in agriculture, as well as in the pharmaceutical and chemical industries. Moreover, a detailed understanding of *B. cinere*a's virulence mechanisms will facilitate the creation of crops with enhanced resistance, reducing dependency on chemical fungicides and contributing to more sustainable and environmentally friendly agricultural systems.

Looking forward, much remains to be discovered about the complex secondary metabolism of *B. cinerea*. Future research prospects include the detailed exploration of newly identified biosynthetic pathways, functional characterization of unknown metabolic products, and investigation of specific molecular interactions between the pathogen and its hosts. Furthermore, advancing our understanding of the genetic and epigenetic regulation of these biosynthetic systems to develop targeted and tailored interventions is essential. Continuing to expand our knowledge in these areas will not only enrich fundamental science in mycology and plant pathology but also have significant practical applications in crop protection and biotechnology.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org. This material includes Figure S1, S2 and S3

Author Contributions: All the authors collaborated in the experiments described in this article. I.S. personally conducted all the experiments. Bioinformatics analyses were conducted jointly by I.S. and C.G. Results and Conclusions were carried out by all authors, I.S., I.G.C. and C.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by a grant from PID-2021-1228990-B-C21 and -C22 funded by MCIN/AEI/ 10.13039/501100011033 and by "ERDF A way of making Europe".

Institutional Review Board Statement Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Please refer to suggested Data Availability Statements in section "MDPI Research Data Policies" at https://www.mdpi.com/ethics.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Farrell, R.E. Transcriptomes and Bioinformatics. In *RNA Methodologies*; Pennsylvania, 2023; pp. 679–696 ISBN 9780323902212.
2. Xu, M.; Xu, C.; Chen, M.; Xiao, Z.; Wang, Y.; Xu, Y.; Xu, D. Comparative Analysis of Commonly Used Bioinformatics Software Based on Omics. *Gene Rep* **2023**, *32*, 101800, doi:10.1016/j.genrep.2023.101800.
3. Li, J.; de Vries, R.P.; Peng, M. Bioinformatics Approaches for Fungal Biotechnology. In *Encyclopedia of Mycology*; Zaragoza, O., Casadevall, A., Eds.; 2021; pp. 536–554 ISBN 9780323851800.
4. Sayers, E.W.; Cavanaugh, M.; Clark, K.; Pruitt, K.D.; Schoch, C.L.; Sherry, S.T.; Karsch-Mizrachi, I. GenBank. *Nucleic Acids Res* **2022**, *50*, doi:10.1093/nar/gkab1135.
5. Ensembl 2023 - PMC Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9825606/ (accessed on 1 March 2024).
6. Kanehisa, M.; Sato, Y. KEGG Mapper for Inferring Cellular Functions from Protein Sequences. *Protein Science* **2020**, *29*, doi:10.1002/pro.3711.
7. Bateman, A.; Martin, M.J.; Orchard, S.; Magrane, M.; Ahmad, S.; Alpi, E.; Bowler-Barnett, E.H.; Britto, R.; Bye-A-Jee, H.; Cukura, A.; et al. UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Res* **2023**, *51*, doi:10.1093/nar/gkac1052.
8. Roth, M.G.; Westrick, N.M.; Baldwin, T.T. Fungal Biotechnology: From Yesterday to Tomorrow. *Frontiers in Fungal Biology* **2023**, *4*, 1135263, doi:10.3389/ffunb.2023.1135263.
9. Bunyat-zada, A.R.; Ross, A.C. Highlights of Bioinformatic Tools and Methods for Validating Bioinformatics Derived Hypotheses for Microbial Natural Products Research. *Curr Opin Chem Biol* **2023**, *76*, 102367, doi:10.1016/j.cbpa.2023.102367.
10. Sarsaiya, S.; Jain, A.; Shi, J.; Chen, J. Fungi Endophytes for Biofactory of Secondary Metabolites: Genomics and Metabolism. In *Biocontrol Agents and Secondary Metabolites: Applications and Immunization for Plant Growth and Protection*; Elsevier Inc., 2020; pp. 1–21 ISBN 9780128229194.

11. Kalkreuter, E.; Pan, G.; Cepeda, A.J.; Shen, B. Targeting Bacterial Genomes for Natural Product Discovery. *Trends Pharmacol Sci* **2020**, *41*, 13–26, doi:10.1016/j.tips.2019.11.002.
12. Amselem, J.; Cuomo, C. a.; van Kan, J. a L.; Viaud, M.; Benito, E.P.; Couloux, A.; Coutinho, P.M.; de Vries, R.P.; Dyer, P.S.; Fillinger, S.; et al. Genomic Analysis of the Necrotrophic Fungal Pathogens Sclerotinia Sclerotiorum and *Botrytis cinerea*. *PLoS Genet* **2011**, *7*, doi:10.1371/journal.pgen.1002230.
13. Collado, I. G. and Viaud, M. Secondary Metabolism in *Botrytis cinerea*: Combining Genomic and Metabolomic Approaches. In *Botrytis – the Fungus, the Pathogen and its Management in Agricultural Systems*; Springer Netherlands, 2016; pp. 291–314.
14. Staats, M.; van Kan, J.A.L. Genome Update of *Botrytis cinerea* Strains B05.10 and T4. *Eukaryot Cell* **2012**, *11*, 1413–1414, doi:10.1128/EC.00164-12.
15. Van Kan, J.A.L.; Stassen, J.H.M.; Mosbach, A.; Van Der Lee, T.A.J.; Faino, L.; Farmer, A.D.; Papasotiriou, D.G.; Zhou, S.; Seidl, M.F.; Cottam, E.; et al. A Gapless Genome Sequence of the Fungus *Botrytis cinerea*. *Mol Plant Pathol* **2017**, *18*, 75–89, doi:10.1111/mpp.12384.
16. Schumacher, J. DHN Melanin Biosynthesis in the Plant Pathogenic Fungus *Botrytis cinerea* Is Based on Two Developmentally Regulated Key Enzyme (PKS)-Encoding Genes. *Mol Microbiol* **2016**, *99*, 729–748, doi:10.1111/mmi.13262.
17. Moraga, J.; Izquierdo-Bueno Reina, I.; Pinedo, C.; Hernández-Galán, R.; Viaud, M.; Collado, I.G. Impairment of Botrydial Production in *Botrytis cinerea* Allows the Isolation of Undescribed Polyketides and Reveals New Insights into the Botcinins Biosynthetic Pathway. *Phytochemistry* **2021**, *183*, 1–7, doi:10.1016/j.phytochem.2020.112627.
18. da Silva Ripardo-Filho, H.; Coca Ruíz, V.; Suárez, I.; Moraga, J.; Aleu, J.; Collado, I.G. From Genes to Molecules, Secondary Metabolism in *Botrytis cinerea*: New Insights into Anamorphic and Teleomorphic Stages. *Plants* **2023**, *12*, 1–31, doi:10.3390/plants12030553.
19. Porquier, A.; Moraga, J.; Morgant, G.; Dalmais, B.; Simon, A.; Sghyer, H.; Collado, I.G.; Viaud, M. Botcinic Acid Biosynthesis in Botrytis Cinerea Relies on a Subtelomeric Gene Cluster Surrounded by Relics of Transposons and Is Regulated by the Zn2Cys6 Transcription Factor BcBoa13. *Current Genetics 2019 65:4* **2019**, *65*, 965–980, doi:10.1007/S00294-019-00952-4.
20. Valero-Jiménez, C.A.; Steentjes, M.B.F.; Slot, J.C.; Shi-Kunne, X.; Scholten, O.E.; van Kan, J.A.L. Dynamics in Secondary Metabolite Gene Clusters in Otherwise Highly Syntenic and Stable Genomes in the Fungal Genus *Botrytis*. *Genome Biol Evol* **2020**, *12*, 2491, doi:10.1093/GBE/EVAA218.
21. Plesken, C.; Pattar, P.; Reiss, B.; Noor, Z.N.; Zhang, L.; Klug, K.; Huettel, B.; Hahn, M. Genetic Diversity of *Botrytis cinerea* Revealed by Multilocus Sequencing, and Identification of *B. cinerea* Populations Showing Genetic Isolation and Distinct Host Adaptation. *Front Plant Sci* **2021**, *12*, doi:10.3389/FPLS.2021.663027/FULL.
22. Singh, M.; Chaudhary, S.; Sareen, D. Non-Ribosomal Peptide Synthetases: Identifying the Cryptic Gene Clusters and Decoding the Natural Product. *J Biosci* **2017**, *42*, 175–187, doi:10.1007/S12038-017-9663-Z.
23. Fernández-Morales, A.; Carbú, M.; González-Rodríguez, V.E.; Papaspyrou, S.; Garrido, C.; Cantoral, J.M. Deletion of the Bcnrps1 Gene Increases the Pathogenicity of *Botrytis cinerea* and Reduces Its Tolerance to the Exogenous Toxic Substances Spermidine and Pyrimethanil. *Journal of Fungi* **2021**, *7*, doi:10.3390/jof7090721.
24. Colmenares, A.J.; Aleu, J.; Durán-Patrón, R.; Collado, I.G.; Hernández-Galán, R. The Putative Role of Botrydial and Related Metabolites in the Infection Mechanism of *Botrytis cinerea*. *J Chem Ecol* **2002**, *28*, 997–1005, doi:10.1023/A:1015209817830/METRICS.
25. Izquierdo-Bueno, I.; González-Rodríguez, V.E.; Simon, A.; Dalmais, B.; Pradier, J.M.; Le Pêcheur, P.; Mercier, A.; Walker, A.S.; Garrido, C.; Collado, I.G.; et al. Biosynthesis of Abscisic Acid in Fungi: Identification of a Sesquiterpene Cyclase as the Key Enzyme in *Botrytis cinerea*. *Environ Microbiol* **2018**, *20*, 2469–2482, doi:10.1111/1462-2920.14258.
26. Suárez, I.; González-Rodríguez, V.E.; Viaud, M.; Garrido, C.; Collado, I.G. Identification of the Sesquiterpene Cyclase Involved in the Biosynthesis of (+)-4-Epi-eremophil-9-en-11-ol Derivatives Isolated from *Botrytis cinerea*. *ACS Chem Biol* **2020**, *15*, 2775–2782, doi:10.1021/acschembio.0c00561.
27. Pinedo, C.; Moraga, J.; Barua, J.; González-Rodríguez, V.E.; Aleu, J.; Durán-Patrón, R.; Macías-Sánchez, A.J.; Hanson, J.R.; Viaud, M.; Hernández-Galán, R.; et al. Chemically Induced Cryptic Sesquiterpenoids and Expression of Sesquiterpene Cyclases in *Botrytis cinerea* Revealed New Sporogenic (+)-4-Epieremophil-9-En-11-Ols. *ACS Chem Biol* **2016**, *11*, 1391–1400, doi:10.1021/acschembio.5b00931.
28. Schumacher, J.; Gautier, A.; Morgant, G.; Studt, L.; Ducrot, P.H.; Le Pêcheur, P.; Azeddine, S.; Fillinger, S.; Leroux, P.; Tudzynski, B.; et al. A Functional Bikaverin Biosynthesis Gene Cluster in Rare Strains of *Botrytis cinerea* Is Positively Controlled by VELVET. *PLoS One* **2013**, *8*, doi:10.1371/JOURNAL.PONE.0053729.
29. Schumacher, J.; Simon, A.; Cohrs, K.C.; Viaud, M.; Tudzynski, P. The Transcription Factor BcLTF1 Regulates Virulence and Light Responses in the Necrotrophic Plant Pathogen Botrytis Cinerea. *PLoS Genet 10* **2014**, *10*, 1–21, doi:10.1371/journal.pgen.1004040.

30. Zhang, C.; He, Y.; Zhu, P.; Chen, L.; Wang, Y.; Ni, B.; Xu, L. Loss of Bcbrn1 and Bcpks13 in Botrytis Cinerea Not Only Blocks Melanization but Also Increases Vegetative Growth and Virulence. *Molecular Plant-Microbe Interactions* **2015**, *28*, 1091–1101, doi:10.1094/MPMI-04-15-0085-R/ASSET/IMAGES/LARGE/MPMI-04-15-0085-R_F8.JPEG.
31. Porquier, A.; Morgant, G.; Moraga, J.; Dalmais, B.; Luyten, I.; Simon, A.; Pradier, J.M.; Amselem, J.; Collado, I.G.; Viaud, M. The Botrydial Biosynthetic Gene Cluster of *Botrytis cinerea* Displays a Bipartite Genomic Structure and Is Positively Regulated by the Putative Zn(II)2Cys6 Transcription Factor BcBot6. *Fungal Genetics and Biology* **2016**, *96*, 33–46, doi:10.1016/j.fgb.2016.10.003.
32. González, M.; Brito, N.; González, C. The *Botrytis cinerea* Elicitor Protein BcIEB1 Interacts with the Tobacco PR5-Family Protein Osmotin and Protects the Fungus against Its Antifungal Activity. *New Phytologist* **2017**, *215*, 397–410, doi:10.1111/NPH.14588.
33. Burkhardt, I.; Ye, Z.; Janevska, S.; Tudzynski, B.; Dickschat, J.S. Biochemical and Mechanistic Characterization of the Fungal Reverse N-1-Dimethylallyltryptophan Synthase DMATS1Ff. *ACS Chem Biol* **2019**, *14*, 2922–2931, doi:10.1021/ACSCHEMBIO.9B00828/SUPPL_FILE/CB9B00828_SI_001.PDF.
34. Schumacher, J.; Simon, A.; Cohrs, K.C.; Traeger, S.; Porquier, A.; Dalmais, B.; Viaud, M.; Tudzynski, B. The VELVET Complex in the Gray Mold Fungus *Botrytis cinerea*: Impact of BcLAE1 on Differentiation, Secondary Metabolism, and Virulence. *Molecular Plant-Microbe Interactions* **2015**, *28*, 659–674, doi:10.1094/MPMI-12-14-0411-R/ASSET/IMAGES/LARGE/MPMI-12-14-0411-R_T3.JPEG.
35. Swaminathan, R.; Huang, Y.; Moosavinasab, S.; Buckley, R.; Bartlett, C.W.; Lin, S.M. A Review on Genomics APIs. **2016**, doi:10.1016/j.csbj.2015.10.004.
36. Seoane, J.A.; Dorado, J.; Aguiar-Pulido, V.; Pazos, A. Data Integration in Genomic Medicine: Trends and Applications Contribution of the IMIA Working Group on Informatics in Genomic Medicine. **2012**.
37. Nilsson, R.H.; Anslan, S.; Bahram, M.; Wurzbacher, C.; Baldrian, P.; Tedersoo, L. Mycobiome Diversity: High-Throughput Sequencing and Identification of Fungi. *Nature Reviews Microbiology 2018 17:2* **2018**, *17*, 95–109, doi:10.1038/s41579-018-0116-y.
38. Ritchie, M.D.; Holzinger, E.R.; Li, R.; Pendergrass, S.A.; Kim, D. Methods of Integrating Data to Uncover Genotype–Phenotype Interactions. *Nature Reviews Genetics 2015 16:2* **2015**, *16*, 85–97, doi:10.1038/nrg3868.
39. Brown, D.W.; Butchko, R.A.E.; Proctor, R.H. Fusarium Genomic Resources: Tools to Limit Crop Diseases and Mycotoxin Contamination. *Mycopathologia* **2006**, *162*, 191–199, doi:10.1007/S11046-006-0053-6/METRICS.
40. Mohanta, T.K.; Al-Harrasi, A. Fungal Genomes: Suffering with Functional Annotation Errors. *IMA Fungus* **2021**, *12*, doi:10.1186/s43008-021-00083-x.
41. Liu, J.; Zhang, Y.; Shi, Y.; Zheng, Y.; Zhu, Y.; Guan, Z.; Shen, D.; Dou, D. FungiExp: A User-Friendly Database and Analysis Platform for Exploring Fungal Gene Expression and Alternative Splicing. *Bioinformatics* **2023**, *39*, doi:10.1093/bioinformatics/btad042.
42. Potgieter, L.; Feurtey, A.; Dutheil, J.Y.; Stukenbrock, E.H. On Variant Discovery in Genomes of Fungal Plant Pathogens. *Front Microbiol* **2020**, *11*, doi:10.3389/fmicb.2020.00626.
43. Iquebal, M.A.; Jaiswal, S.; Mishra, V.K.; Jasrotia, R.S.; Angadi, U.B.; Singh, B.P.; Passari, A.K.; Deka, P.; Prabha, R.; Singh, D.P.; et al. Fungal Genomic Resources for Strain Identification and Diversity Analysis of 1900 Fungal Species. *Journal of Fungi* **2021**, *7*, doi:10.3390/jof7040288.
44. Basenko, E.Y.; Pulman, J.A.; Shanmugasundram, A.; Harb, O.S.; Crouch, K.; Starns, D.; Warrenfeltz, S.; Aurrecoechea, C.; Stoeckert, C.J.; Kissinger, J.C.; et al. FungiDB: An Integrated Bioinformatic Resource for Fungi and Oomycetes. *Journal of Fungi* **2018**, *4*, 1–28, doi:10.3390/jof4010039.
45. Pérez-Lara, G.; Moyano, T.C.; Vega, A.; Larrondo, L.F.; Polanco, R.; Álvarez, J.M.; Aguayo, D.; Canessa, P. The *Botrytis cinerea* Gene Expression Browser. *Journal of Fungi* **2023**, *9*, doi:10.3390/JOF9010084.
46. Zhang, N.; Yang, Z.; Liang, W.; Liu, M. Global Proteomic Analysis of Lysine Crotonylation in the Plant Pathogen *Botrytis cinerea*. *Front Microbiol* **2020**, *11*, doi:10.3389/FMICB.2020.564350.
47. Xu, Y.; Li, X.; Liang, W.; Liu, M. Proteome-Wide Analysis of Lysine 2-Hydroxyisobutyrylation in the Phytopathogenic Fungus *Botrytis cinerea*. *Front Microbiol* **2020**, *11*, doi:10.3389/FMICB.2020.585614.