BMC Genomics

RESEARCH

Open Access

Six type-I PKS classes and highly conserved melanin and elsinochrome gene clusters found in diverse *Elsinoë* species

Alishia van Heerden[1], Nam Q. Pham[1], Brenda D. Wingfield[1], Michael J. Wingfield[1] and P. Markus Wilken[1*]

Abstract

Elsinoë species are phytopathogenic fungi that cause serious scab diseases on economically important plants. The disease symptoms arise from the effects of a group of phytotoxins known as elsinochromes, produced via a type-I polyketide synthase (PKS) biosynthetic pathway. The elsinochrome gene cluster was first annotated in *Elsinoë fawcettii* where the main type-I PKS gene was characterized as *EfPKS1*. A later study showed that this gene and the associated cluster had not been correctly annotated, and that *EfPKS1* was actually the anchor gene of the melanin biosynthetic pathway. A new type-I PKS gene *EfETB1* associated with elsinochrome production was also identified. The aim of this study was to identify all type-I PKS genes in the genomes of seven *Elsinoë* species with the goal of independently verifying the PKS containing clusters for both melanin and elsinochrome production. A total of six type-I PKS classes were identified, although there was variation between the species in the number and type of classes present. Genes similar to the *E. fawcettii EfPKS1* and *EfETB1* type-I PKS genes were associated with melanin and elsinochrome production respectively in all species. The complete melanin and elsinochrome PKS containing clusters were subsequently annotated in all the species with high levels of synteny across *Elsinoë* species. This study provides a genus-level overview of type-I PKS distribution in *Elsinoë* species, including an additional line of support for the annotation of the melanin and elsinochrome PKS containing clusters in these important plant pathogens.

Keywords Perylenequinone, Phylogenetic analysis, Polyketide synthase, Secondary metabolites, Synteny analysis

Introduction

Species of *Elsinoë* (Elsinoaceae, Myriangiales) are specialized phytopathogenic fungi that cause serious shoot and foliar diseases on many plant species, including agronomic crops and trees of economic importance [1]. Well-known species include *E. fawcettii* and *E. australis* causing citrus scab [2], *E. perseae* that causes avocado scab [3], *E. ampelina* the cause of grapevine spot anthracnose [4], and the recently described *E. necatrix* and *E. masingae* that result in a devastating Eucalyptus scab and shoot malformation disease in Indonesia [5] and South Africa [6]. *Elsinoë* infections are characterized by cork-like, cracked lesions known as scabs that can often lead to distortion and necrosis of the infected tissues [1, 5, 7]. These disease symptoms are initiated by a group of secondary metabolites known as elsinochromes [8, 9].

Elsinochromes are well-known secondary metabolites found in *Elsinoë* spp. that are crucial components of fungal virulence and lesion formation [9–11]. The red and orange pigmented elsinochromes are described as photosynthesizing agents as they produce reactive oxygen species in aerobic and light activated conditions [9]. These reactive oxygen species damage cell membranes and

*Correspondence:
P. Markus Wilken
markus.wilken@fabi.up.ac.za
[1] Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa

cause electrolyte leakage, which ultimately lead to tissue necrosis [9, 12]. Elsinochromes are structurally similar to perylenequinone phytotoxins synthesized by many other fungi, including aspergillin produced by *Aspergillus niger* [13], hypomycin produced by *Hypomyces* spp. [14], and cercosporin produced by *Cercospora* spp. [12].

Melanin is a well-known secondary metabolite produced by a large number of fungi where it protects against multiple environmental agents including UV radiation, enzymatic lysis and oxidative damage [15]. Melanins are classified according to the composition of the pigment [16], resulting in four types, the eumelanins, pheomelanins, DHN-melanins, and pyomelanins [17]. Similar to perylenequinones [18], synthesis of DHN-melanin is regulated by a polyketide synthase that catalyzes the first step in the pigment production pathway [17]. These multi-functional enzymes known as polyketide synthase (PKS), encoded by co-regulated genes, are responsible for the structural and functional diversity of secondary metabolites such as melanin [19].

The type-I polyketide synthase (PKS) gene thought to be responsible for elsinochrome production was first characterized in *E. fawcettii* [10]. The authors cloned a type-I PKS gene that they named *EfPKS1* and assigned this gene as the primary PKS responsible for elsinochrome production. Their findings were substantiated by a phylogenetic analysis, revealing that this gene was similar to other polyketide synthase (PKS) genes known to be involved in pigment production [20]. Targeted disruption of *EfPKS1* produced mutants that showed lower levels of virulence and were unable to produce detectable levels of elsinochrome. In a subsequent study, several other genes in the vicinity of *EfPKS1* were shown to form part of the elsinochrome cluster [20].

A study by Ebert et al. [21] refuted the findings of Chung and Liao [20], rather linking the *EfPKS1* gene to melanin biosynthesis in *E. fawcettii*. The authors also showed that *EfPKS1* was most closely related to type-I PKS genes responsible for melanin production in other fungi [21]. This was experimentally substantiated by a gene knock-out that showed the Δ*EfPKS1* mutant produced pale fungal colonies with lower levels of melanin, but left elsinochrome production unchanged. Using a draft genome sequence, Ebert et al. [21] described an amended gene cluster responsible for elsinochrome production, identifying *EfETB1* as the core type-I PKS gene in this cluster. A mutant culture in which *EfETB1* was disrupted lost the ability to produce elsinochrome, although melanin production was not altered [21]. These findings supported the role of *EfPKS1* and the associated gene cluster in the melanin biosynthesis pathway, while the *EfETB1* cluster was likely involved in elsinochrome production [21].

The studies of Chung and Liao [20] and Ebert et al. [21] provided experimental and phylogenetic support for their findings, although these studies produced contrasting results. Both studies included only a single *Elsinoë* species for the bioinformatic analysis, but the recent availability of genome sequences for several economically important *Elsinoë* species [21–25] provide the opportunity to further evaluate the pathways involved in melanin and elsinochrome production in these fungi.

A recent study by Cequeña and Sumabat-Dacones [26] provided an *in-silico* analysis of putative effectors and secondary metabolite synthetic clusters present in six *Elsinoë* genomes sequenced prior to 2021. Although the study provided insight into the potential secondary metabolite clusters present in different *Elsinoë* species, including that of melanin and elsinochrome biosynthesis, it did not address the missanotation of the elsinochrome cluster by Chung and Liao [20] as reported by Ebert et al. [21]. Furthermore, the structure of the large elsinochrome cluster was only annotated based on the output of a single annotation pipeline, with no further validation.

In the present study, a phylogenetics-based approach was used to independently identify and compare the type-I PKS genes present in the genomes of seven *Elsinoë* species, including the recently sequenced genome of *E. necatrix* [24]. For comparative purposes, the type-I PKS genes in two closely related Dothideomycetes, *Myriangium duriaei* and *Aureobasidium pullulans*, were also analysed. The results were used to identify the type-I PKS containing clusters thought to be involved in melanin and perylenequinone production across all species included in the study and to compare these with those from the previous studies by Chung and Liao [20] and Ebert et al. [21].

Materials and methods

Identification and characterization of type-I PKS genes

Available whole genome sequences for seven *Elsinoë* (*E. ampelina, E. arachidis, E. australis, E. batatas, E. fawcettii, E. murrayae* and *E. necatrix*) and two related species, *Myriangium duriaei* (Myriangiaceae, Myriangiales) and *Aureobasidium pullulans* (Saccotheciaceae, Dothideales), were retrieved from the public assembly database at NCBI [27] (Table 1). Basic genome statistics and genome completeness for all nine selected genomes were assessed on the French Galaxy platform (usegalaxy.fr) [28] using QUAST [29] and BUSCO [30], respectively (Table S1). The genome sequences were then annotated using GeneMark-ES, with parameters optimised for fungi [31], and imported into Geneious Prime 2021.0.3 [32] to generate a genomic database. This database was mined for type-I PKS genes by predicting secondary metabolite (SM) clusters using the fungal version of antiSMASH [33].

Table 1 Genome sequences used in this study

Species	Strain	Accession number	Reference
Elsinoë ampelina	YL-1	SMYM00000000	Li et al. [35]
Elsinoë arachidis	LNFT-H01	JAAPAX010000000	Jiao et al. [23]
Elsinoë australis	Hillstone_3	QGIG00000000	NCBI – BioProject PRJNA419078
Elsinoë batatas	CRI-CJ2	JAESVG020000000	Zhang et al. [36]
Elsinoë fawcettii	DAR-70024	SWCR00000000	Shanmugam et al. [37]
Elsinoë murrayae	CQ-2017a	NKHZ00000000	NCBI – BioProject PRJNA385043
Elsinoë necatrix	CMW 56129	JANZYI000000000	Pham et al. [25]
Myriangium duriaei	CBS 260.36	JAAEIR000000000	Haridas et al. [38]
Aureobasidium pullulans	EXF-150	AYEO00000000	Gostinčar et al. [39]

For each of the type-I PKS SM clusters predicted by antiSMASH, the main type-I PKS gene was mapped back to the respective genome sequence using the genomic position provided by antiSMASH (Table 2). The respective genes were translated to proteins in Geneious Prime and submitted to the BLASTp server [34] at NCBI to confirm their identity.

To determine the putative function of the different PKS proteins, a phylogenetic approach was used. For the purpose of the study, the broad term melanins was used to refer to the DHN-type melanin synthesized by a polyketide synthase. First, reference protein sequences for eight known PKS groups, i.e., melanins, perylenequinones, macrolides, t-toxins, aflatoxins, azaphilones, anthraquinones, and naphthoquinones, were retrieved from the protein database on NCBI [34] (Table 3). Protein domains characteristic of PKS proteins were identified from all available PKS sequences using InterProScan [40]. The protein sequences were trimmed to retain only the KS and AT domains, and the resulting sequences were aligned using MAFFT v. 7 [41]. The veracity of the alignment was visually confirmed in MEGA v. 7 [42]. Maximum Likelihood (ML) analysis was conducted using RaxML v. 8.2.4 [43] on the CIPRES Science Gateway v. 3.3 [44] with a default GTR substitution matrix and 1,000 rapid bootstraps. Resulting phylogenetic tree was viewed and edited in MEGA v. 7 [42].

Annotation and synteny analysis

Cluster boundaries provided by the antiSMASH analysis were used to define the putative melanin and perylenequinone PKS containing clusters from the *Elsinoë* species and the two related taxa (*M. duriaei* and *A. pullulans*). These boundaries were then trimmed to align with those reported previously for *E. fawcettii* [20, 21]. Each of the genes within the newly defined cluster were translated to amino acid sequences using Geneious Prime and submitted to the BLASTp server at NCBI for confirmation of protein identity (Table S2).

The two gene clusters were extracted from the genomes of all the *Elsinoë* species and the two closely related taxa for a synteny analysis. Clinker 0.0.21 [45] was used to assess global alignment between the cluster regions, with the optimal display order determined by cluster similarity. An interactive visualisation of the synteny was also generated in Clinker using clustermap.js using default settings [45]. Genes were grouped based on cluster similarity as determined by Clinker and were labelled according to their assigned protein identity previously determined by BLASTp on NCBI (Fig. 3).

Results

Identification and characterization of type-I PKS genes

AntiSMASH analysis of the *Elsinoë* genomes revealed regions representing type-I polyketide synthase (PKS), non-ribosomal peptide synthetase (NRPS), terpenes and fungal-RiPP like secondary metabolite clusters (Table S3). Seven type-I PKS clusters were found in *E. necatrix*, five in *E. arachidis*, *E. batatas* and *A. pullulans*, four in *E. fawcettii* and *M. duriaei* and three in *E. ampelina*, *E. murrayae* and *E. australis* (Table 2). Each type-I PKS cluster had a single main type-I PKS gene which was mapped back to the respective genomes and for which gene identity was confirmed by BLASTp analysis. A phylogenetic comparison of the highly conserved ketosynthase (KS) and acyltransferase (AT) protein domains from the predicted type-I PKS genes made it possible to assign a putative function to each type-I PKS gene based on a set of reference proteins [21, 23, 46] (Table 3). Eight distinct classes of PKSs were identified, with all members within a class noted to synthesize structurally similar secondary metabolites (Fig. 1). Based on the clustering of the identified PKS protein sequences and the reference protein sequences within the tree, the classes were categorized as perylenequinones, aflatoxins,

Table 2 Type-I PKS genes identified by antiSMASH, mapped back to the *Elsinoë* genome sequences. Predicted PKS class determined by phylogenetic comparison

	Genome position			
Species	**Contig number**	**Nucleotide position**	**Gene number**	**Predicted PKS family**
Elsinoë necatrix	tig00000013	784,721–791,348	3489	Melanin
	tig00000013	1,558,466–1,565,254	3763	Elsinochrome
	tig00000019	429,679–437,037	5192	T-toxin 2
	tig00000027	8,672–15,015	5917	Macrolides
	tig00000031	26,392–32,738	6760	Macrolides
	tig00000032	2,289,787–2,297,447	7648	T-toxin 1
	tig00000032	18,922–26,080	6764	T-toxin 3
Elsinoë fawcettii	SWCR01000008.1	125,267–132,055	8388	Elsinochrome
	SWCR01000008.1	926,368—932,998	8689	Melanin
	SWCR01000009.1	1,090,834—1,096,384	7017	Anthraquinone
	SWCR01000004.1	1,086,308—1,095,671	2761	T-toxin 5
Elsinoë ampelina	SMYMO100002.1	2,252,206–2,258,822	3531	Melanin
	SMYMO100002.1	3,139,385–3,146,179	3828	Elsinochrome
	SMYMO100002.1	1,622,720–1,629,309	3315	Naphthoquinone
Elsinoë arachidis	JAAPAX010000003.1	3,536,380–3,543,171	4012	Elsinochrome
	JAAPAX010000003.1	2,540,881–2,547,495	3699	Melanin
	JAAPAX010000003.1	1,905,809–1,912,398	3479	Naphthoquinone
	JAAPAX010000002.1	38,409–45,945	1563	T-toxin 3
	JAAPAX010000006.1	2,932,477–2,940,121	7382	T-toxin 1
Elsinoë batatas	JAESVG020000002.1	2,613,668–562,672	2396	Elsinochrome
	JAESVG020000002.1	1,315,266–1,321,882	2108	Melanin
	JAESVG020000002.1	1,805,220–1,810,667	1899	Naphthoquinone
	JAESVG020000005.1	2,090,919–2,100,402	5267	T-toxin 4
	JAESVG020000010.1	3,540,544–3,548,613	9482	T-toxin 1
Elsinoë murrayae	NKHZ01000011.1	83,358–90,130	4959	Elsinochrome
	NKHZ01000011.1	738,183–744,800	5223	Melanin
	NKHZ01000086.1	323,643–331,347	5484	T-toxin 1
Elsinoë australis	SWCS01000013.1	191,675–185,030	3198	Melanin
	SWCS01000013.1	996,526–1,003,321	3516	Elsinochrome
	SWCS01000010.1	1,256,623–1,262,193	8180	Anthraquinone
Myriangium duriaei	ML996083.1	1,696,308–1,702,980	3223	
	ML996081.1	3,101,374–3,108,826	1166	
	ML996086.1	1,365,136–1,357,142	5558	
	ML996088.1	874,394–866,632	6568	
Aureobasidium pullulans	KL584978.1	128,457–136,022	3177	
	KL584984.1	845,879–852,633	6316	
	KL584985.1	363,453–370,232	6540	
	KL584986.1	318,552–325,043	6838	
	KL585017.1	4,394–13,826	10,666	

naphthoquinones, melanins, anthraquinones, macrolides, t-toxins and azaphilones [21, 23, 46].

Only the melanin and perylenequinone gene clusters were represented in all the *Elsinoë* species considered in this study. The *Elsinoë* PKS protein domains present in the melanin class were representative of the melanin biosynthesis pathway, while the proteins clustering in the perylenequinones class represented the elsinochrome producing PKS gene cluster previously identified by Ebert et al. [21]. Putative amino acid sequences from specific *Elsinoë* species were found to represent the classes of macrolides, naphthoquinones, t-toxins,

Table 3 Reference sequences for eight known PKS gene classes

PKS gene class	Pathogen	Protein accession number
Melanins	*Colletotrichum lagenarium*	BAA18956
	Nodulisporium sp.	AAD38786
	Sordaria macrospore	CAM35471
Perylenequinones	*Cercospora nicotianae*	AAT69682.1
	Cercospora zeina	ARU80380
	Cercospora beticola	XP023460065
	Mycosphaerella coffeicola	ADO14890
Anthraquinones	*Aspergillus fumigatus*	XP_746435.1
	Parastagonospora nodorum	EAT76667.2
	Aspergillus nidulans	XP_657754.1
	Aspergillus terreus	XP_001217072.1
T-toxins	*Bipolaris maydis*	N4WHA7
	Bipolaris maydis	*AAB08104*
	Bipolaris maydis	N4WHA3
Aflatoxins	*Aspergillus flavis*	AAS90093.1
	Dothistroma septosporum	EME39092.1
	Aspergillus nidulans	AAC49191.1
Naphthoquinones	*Fusarium fujikuroi*	CAB92399.1
	Fusarium graminearum	AAU10633.1
	Aspergillus nidulans	CAA46695.2
	Aspergillus niger	EHA28527.1
Macrolides	*Hypomyces subiculosus*	ACD39762.1
	Fusarium graminearum	ABB90282.1
Azaphilones	*Monascus purpureus*	BAD44749.1
	Penicillium brevicompactum	ADY00130.1

and anthraquinones. All *Elsinoë* species other than *E. ampelina* and *E. australis* had one or more putative PKS proteins present in the t-toxin class, with *E. necatrix* having three representatives. *E. australis* and *E. fawcettii* had one putative PKS protein each in the anthraquinones class, while *E. ampelina, E. arachidis* and *E. batatas* each had a single putative PKS protein in the naphthoquinones class. *E. necatrix* was the only species that had any representation (two PKS proteins) in the macrolides class. In addition, none of the *Elsinoë* species had any representative putative protein sequence in the aflatoxins class. *M. duriaei* and *A. pullulans* included for comparative purposes had representative putative protein sequences only in three classes. Both *A. pullulans* and *M. duriaei* had a single PKS gene in the melanins class, while four *A. pullulans* and one *M. duriaei* PKS genes were present in the t-toxins group. Two PKS genes from the genome of *M. duriaei* were present in the azaphilones outgroup, with *A. pullulans* not having a representative in this group.

InterProScan analysis of the different *Elsinoë* putative PKS proteins identified eleven different domains (Fig. 2). All of the identified PKSs contained the three main domains i.e., β-ketosynthase (KS; Pfam accession number PF00109), acyl-transferase (AT; PF00698) and ACP (PP-binding; PF00550). The melanin, elsinochrome and macrolides PKSs were very similar with all three containing the additional starter unit (SAT; PF16073), dehydratase (DH; PF14765) and thioesterase (TE; PF00975) domains. Similarly, the naphthoquinones PKSs also contained an additional SAT and TE domain but lacked a DH domain. The t-toxin PKSs were divided into five groups depending on the additional domains that were present. T-toxin 1 PKS was defined by the DH, methyl-transferase (MT; PF08242), trans-acting enoyl (ER; PF00107) and β-ketoreductase (KR; PF08659) domains. T-toxin 2 PKSs contained the DH, MT, and ER domains, while t-toxin 3 contained DH, ER and KR domains. T-toxin 4 PKS contained a DH, MT, KR and a nonribosomal peptide synthetase (NRPS; PF00668) domain, while t-toxin 5 PKS contained an acetyl-CoA synthetase (AMP; PF00501) domain in addition to the DH, MT, and KR domains.

Annotation and synteny analysis of the elsinochrome and melanin biosynthetic gene cluster

The melanin and perylenequinone type-I PKS genes were used to annotate the complete melanin and elsinochrome gene clusters in all seven *Elsinoë* species. After adjusting the cluster boundaries to match those previously described [20, 21], the melanin and elsinochrome clusters contained six and ten genes, respectively. BLASTp using the putative cluster proteins as queries confirmed that each gene matched to a homolog previously characterised in *E. fawcettii* (Table S3) [20, 21]. For the melanin biosynthetic cluster, this included the core biosynthetic PKS gene encoding a polyketide synthase (*EfPKS1)*, as well as genes for transcription factor

(See figure on next page.)

Fig. 1 Phylogeny of PKSs of related *Elsinoë* species and reference species. Maximum likelihood phylogenetic tree illustrating the phylogenetic relationship of all predicted non-reducing polyketide synthases (PKSs) from the selected species set (Table 3) plus those derived from the set of PKSs in *Elsinoë* species determined by AntiSMASH. The tree was constructed from the aligned full-length PKS-AT and PKS-KS domains. Established biosynthetic end products for PKSs is indicated by the background colour, highlighting Melanins, Naphthoquinones, Anthraquinones, Perylenequinones, Aflatoxins, T-toxins, Macrolides and the outgroup class, Azaphilones. The values on the tree nodes represent the percentage bootstrap support

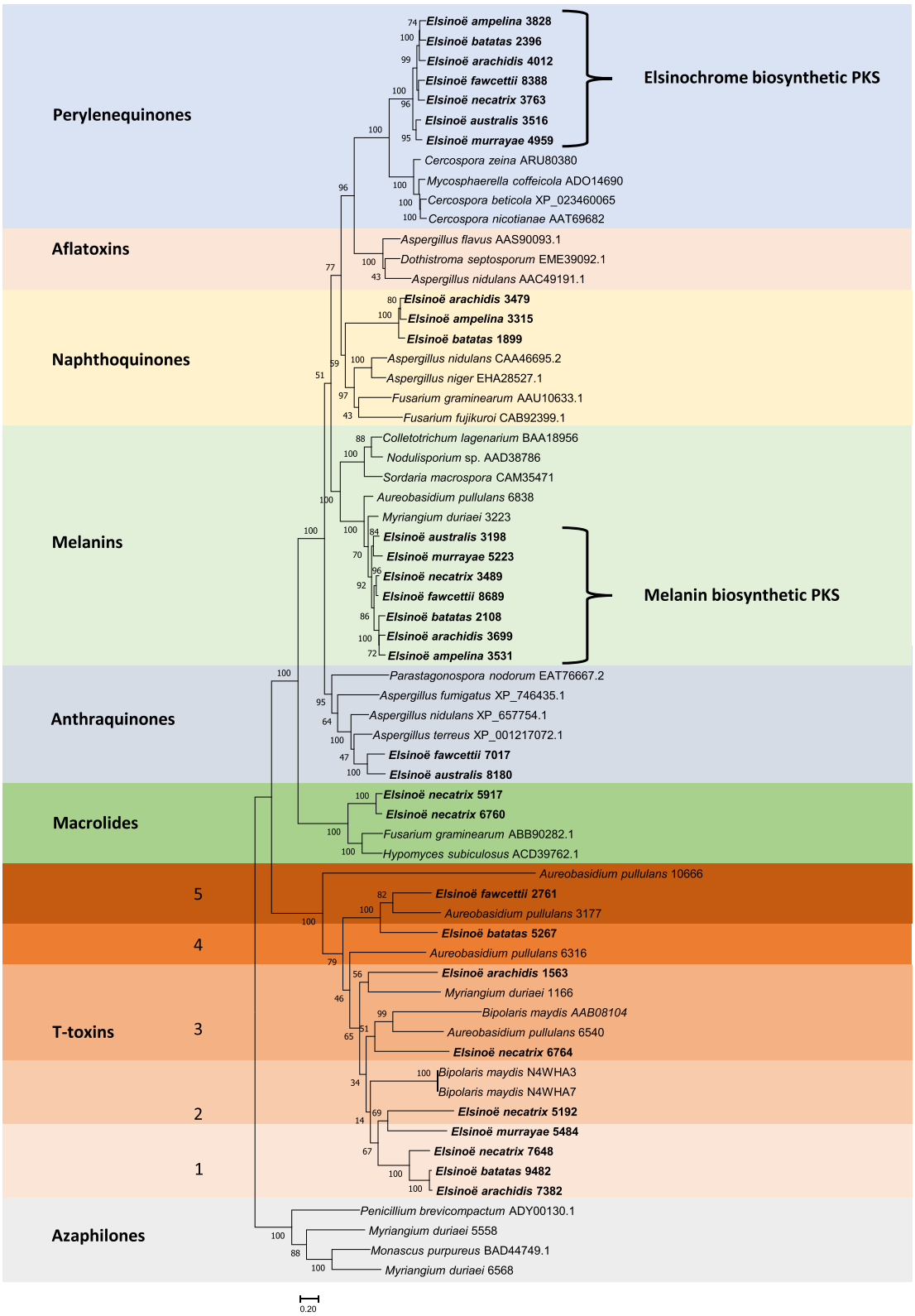

Fig. 1 (See legend on previous page.)

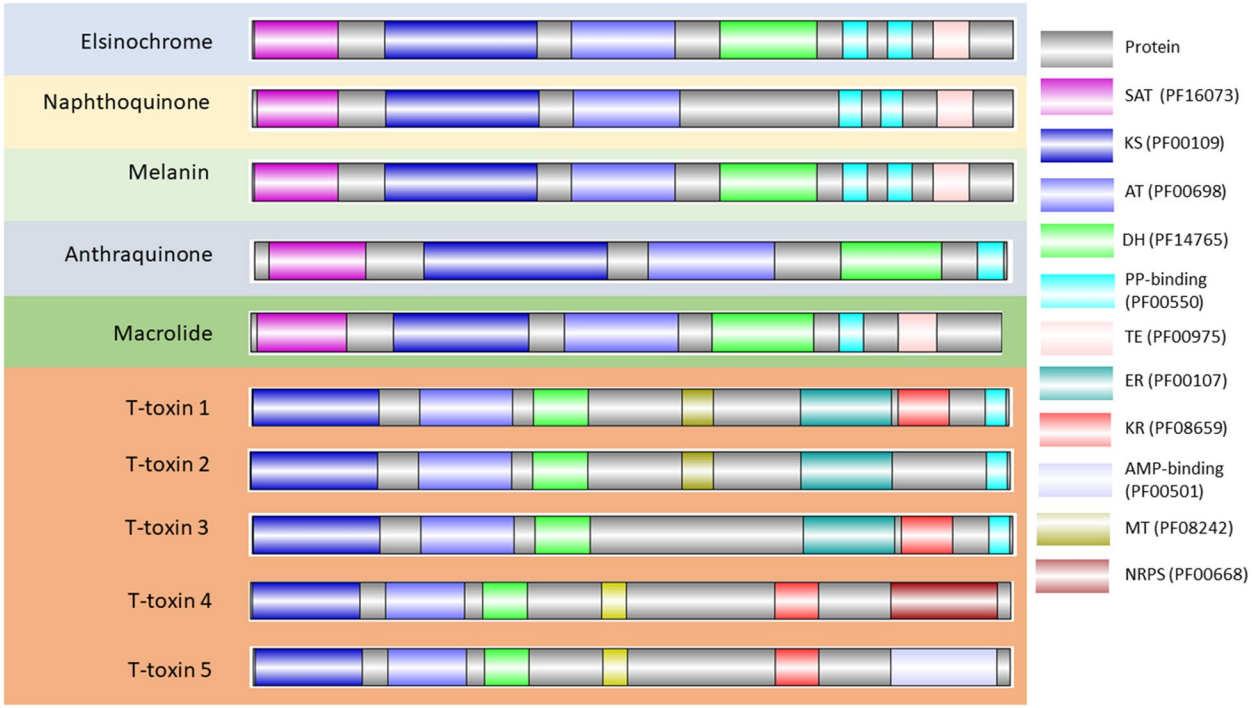

Fig. 2 Structure of polyketide synthase proteins. The conservative domain of polyketide synthases was clarified by InterProScan, and the visualization of different domains by using DOG 2.0 [47]. KS = β-ketosynthase (Pfam accession number_PF00109), AT = acyl-transferase (PF00698), PP-binding = ACP like protein (PF00550), SAT = Starter unit (PF16073), DH = Dehydratase (PF14765), TE = Thioesterase (PF00975), MT = Methyl-transferase (PF08242), ER = trans-acting enoyl (PF00107), KR = β-ketoreductase (PF08659), NRPS = Nonribosomal peptide synthetase (PF00668) AMP-binding = Acetyl-CoA synthetase (PF00501)

(*TSF1*), ESC reductase (*RDT1*), ESC prefoldin protein subunit 3 (*PRF1*), and ECT1 transporter (*ECT1*). For the elsinochrome biosynthetic cluster, the predicted genes included the core biosynthetic PKS gene encoding the protein polyketide synthase CTB1-like protein (*CTB1*) as well as homologs to the genes for O-methyltransferase (*OMT1*), FAD-binding domain-containing protein (*FAD1*), major facilitator superfamily domain-containing protein (*MFS1*), O-methyltransferase 2 (*OMT2*), fungal Zn(2)-Cys(6) binuclear cluster domain-containing protein (ZNC1), protein STB3 (STB3), FAD binding/ oxidoreductase (*OXR1*), and a fasciclin domain-containing protein (*FAS1*).

Only a melanin biosynthesis cluster was detected in the genomes of the closely related Dothideomycetes taxa, *M. duriaei* and *A. pullulans*. In *M. duriaei*, all five genes in the melanin cluster could be identified, while only four of these genes (no BLAST identity to the ECT1 protein sequence) were present in *A. pullulans*. No homologs to any of the genes from the elsinochrome biosynthetic cluster were detected in the genome sequences of either of the two fungi.

High levels of synteny were apparent within both the melanin and elsinochrome biosynthetic clusters across the different *Elsinoë* species (Fig. 3). For the melanin cluster, the five main genes were present in the order *RDT1—TSF1—PKS1—PRF1—ECT1*. The only exception was the *ECT1* gene that was inverted and positioned upstream of *RDT1* in *E. australis* and *E. murrayae*, resulting in a gene order of *ECT1—RDT1—TSF1—PKS1—PRF1*. Additional genes were also identified between the inverted *ECT1* gene and the *RDT1* gene in these two species. The core melanin cluster genes in *M. duriaei* and *A. pullulans* had the same order and orientation as the consensus for the *Elsinoë* species (*RDT1—TSF1—PKS1—PRF1—ECT1*), apart from the *ECT1* gene that was absent in *A. pullulans*. Similarly, the structure of the elsinochrome cluster was highly conserved between the different *Elsinoë* species. In all the species, the genes were present in the order *CTB1 – OMT1 – FAD1 – MFS1 – OMT2 – ZNC1 – STB3 – FAD2 – OXR1 – FAS1*. A gene encoding for a hypothetical protein, lacking a conserved domain, was found between *MFS1* and *OMT2* in the genomes of *E. batatas* and *E. murrayae*.

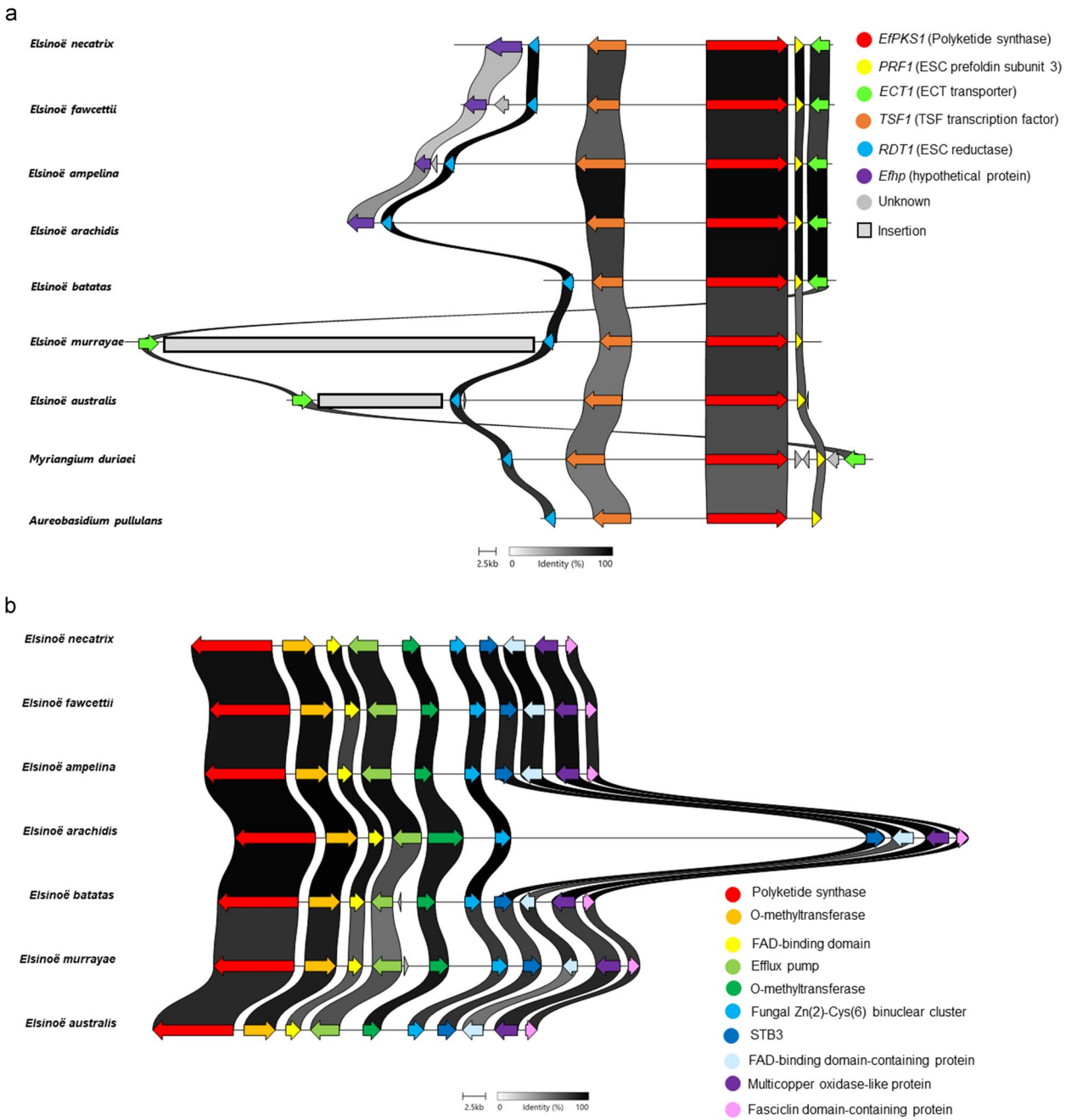

Fig. 3 Synteny analysis map generated in Clinker 0.0.21 of the A_Melanin gene cluster and B_Elsinochrome gene cluster found in seven *Elsinoë* species and two closely related species where applicable. Arrows indicate the orientation of the coding gene and specific genes are shown in a colour bar. The grey bar in figure A indicates an insertion located in the toxin cluster of *E. murrayae* and *E. australis*. The lines connecting the arrows represent gene-encoding proteins that share sequence identity more significant than 40%

Discussion

In this study, an extensive analysis of the type-I PKS genes present in the genomes of seven different *Elsinoë* species was conducted. This analysis not only provided an inventory of the PKS secondary metabolite clusters present in these species, but also identified the complete melanin and elsinochrome PKS containing clusters across all species. A high diversity of type-I PKSs was found in these *Elsinoë* genomes, with representatives of the main type-I PKS genes grouping in six of the eight known PKS classes analysed. Among these, only gene sequences from two classes were present in all seven *Elsinoë* species. These

two classes included the perylenequinones type-I PKS class, representing the elsinochrome biosynthetic cluster, and the melanin type-I PKS class. Further analysis of these two clusters showed strong similarity to the *E. fawcettii* elsinochrome and melanin biosynthetic clusters previously described by Ebert et al. [21]. This study provides a robust comparison of type-I PKSs between the different *Elsinoë* species, and presents a genus level comparison and analysis of both the melanin and elsinochrome gene clusters in these fungi.

A phylogenetic analysis of all the identified type-I PKS genes across the genomes of the seven *Elsinoë* species identified these as responsible for melanin, perylenequinone, anthraquinone, t-toxin, naphthoquinone and macrolide production. Although functional studies would be required to elucidate the role of each of these clusters in the biology of these fungi, previous studies provide some clues. Quinones such as naphthoquinones, perylenequinones or anthraquinones can have multiple functional roles in fungi [48] including acting as antibacterial agents [49], allelochemicals that kill or inhibit competing organisms [50], or photosensitizers that produce reactive oxygen species to damage target plants [51]. Macrolides have also been associated with important fungal biological activities such as cytotoxicity, and can act as antiviral, antibacterial, anti-plasmodial or anti-inflammatory agents [52–54]. Finally, t-toxins are described as host-selective toxins that assist in plant–fungus interactions [55]. The functions of these secondary metabolites are generally governed by the ecological niche in which the species occur [56]. It would consequently be interesting to determine the function of these secondary metabolites in the different *Elsinoë* species.

The two type-I PKS clusters present in all seven *Elsinoë* species were identified as the melanin and perylenequinone (presented as elsinochrome in *Elsinoë* species) type-I PKS clusters. These findings are consistent with those of Cequeña and Sumabat-Dacones [26] where a PKS gene was identified for both melanin and elsinochrome production in the genomes of the six *Elsinoë* species. Elsinochrome is an important polyketide secondary metabolite produced exclusively by *Elsinoë* species [9, 11, 57], while melanin is a pigment present in a wide range of fungi [15]. Although both the elsinochrome and melanin PKS containing clusters were present in all seven *Elsinoë* genomes, only the latter was also present in the two non-*Elsinoë* *taxa* included for comparative purposes. These findings provide independent support for the clusters assigned by Ebert et al. [21], and together with the functional analyses provided by those authors, identifies the elsinochrome gene cluster with confidence. It would be interesting in the future, to explore the findings of Liao and Chung [10] suggesting that deletion of the PKS type-I gene responsible for melanin production resulted in lower levels of pathogenicity and elsinochrome production.

In the current study, both the elsinchrome cluster and the melanin cluster were conserved in all seven *Elsinoë spp.* with high levels of synteny between the different species. SM clusters often evolve in specific fungal linages as a response to unique ecological requirements and they tend to remain as a cluster to ensure functionality [58]. The structure and conservation of the elucidated clusters in the present study provides evidence that SM clusters are maintained in closely related species by factors such as natural selection, epistatic selection, or selection for co-regulation through chromatin remodelling [59–63] or horizontal gene transfer [58, 64, 65]. The maintenance of these SM clusters even indicates that it could play a role in the pathogenicity of these fungal pathogens.

Characterization of different PKS clusters in this study relied on robust genomic evidence and accurate identification of the PKS clusters from previous studies [21, 23, 46]. Our *in-silico* approach offers a glimpse into the potential roles of these polyketide biosynthesis clusters, although such an analysis has inherent limitations. Detailed functional studies on these predicted PKS clusters, such as the those by Chung and Liao [20] and Ebert et al. [21] combining gene expression data and analytical chemistry, will provide more substantial support for the results of the current study. Future work could include determining the functional role of these different secondary metabolites in *Elsinoë* species through gene knockout approaches that might in turn provide a key to disease control measures [66]. Nevertheless, this study still provides a comprehensive resource with extensive analysis of the type-I PKS genes for future investigation into the evolution of type-I PKS genes and the function of these clusters in the biology of an economically important, but relatively understudied group of fungi.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-10920-z.

Supplementary Material 1. Genome statistics of the seven *Elsinoë* genomes and two closely related species. The genome statistics and genome completeness of the nine genomes were evaluated using Galaxy France (https://usegalaxy.fr/) to determine the N50 and L50 counts using Quast as well as the BUSCO using the fungal database as a reference.

Supplementary Material 2. BLASTp search of the translated protein sequences against the GenBank database of the *Elsinoë* Melanin and Elsinchrome PKS genes.

Supplementary Material 3. Types of biosynthetic gene clusters predicted by AntiSMASH in each of the nine selected species.

Supplementary Material 4. Melanin gene cluster. Annotation of the melanin gene cluster of all seven *Elsinoë* species and two closely related species used in this study (GenBank format).

Supplementary Material 5. Elsinochrome gene cluster. Annotation of the elsinochrome gene cluster of all seven *Elsinoë* species used in this study (GenBank format).

Acknowledgements
This study was initiated through the bilateral agreement between the Forestry and Agricultural Biotechnology Institute (FABI) and the April Group, RGE, Indonesia.

Authors' contributions
Leadership and funding for this study was provided by P.M.W., B.D.W. and M.J.W. with conceptulisation and design contributed by all authors. N.Q.P annotated the genomes and A.V.H. performed all bioinformatic analysis. All authors contributed to the interpretation of the data and A.V.H. prepared the initial manuscript and all authors contributed to its completion.

Funding
Funding for this study was provided by the National Research Foundation, South Africa and the University of Pretoria.

Data Availability
The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 23 July 2024 Accepted: 18 October 2024
Published online: 22 October 2024

References

1. Fan XL, Barreto RW, Groenewald JZ, Bezerra JD, Pereira OL, Cheewangkoon R, Mostert L, Tian CM, Crous PW. Phylogeny and taxonomy of the scab and spot anthracnose fungus *Elsinoë* (Myriangiales, Dothideomycetes). Stud Mycol. 2017;87:1–41.
2. Chung KR. *Elsinoë fawcettii* and *Elsinoë australis*: the fungal pathogens causing citrus scab. Mol Plant Pathol. 2011;12(2):123–35.
3. Jenkins AE. *Sphaceloma perseae*, the cause of avocado scab. J Agric Res. 1934;49:859–69.
4. Magarey RD, Emmett RW, Magarey PA, Franz PR. Evaluation of control of grapevine anthracnose caused by *Elsinoë ampelina* by pre-infection fungicides. Australas Plant Pathol. 1993;22(2):48–52.
5. Pham NQ, Marincowitz S, Solís M, Duong TA, Wingfield BD, Barnes I, Slippers B, Muro Abad JI, Durán A, Wingfield MJ. Eucalyptus scab and shoot malformation: a new and serious foliar disease of *Eucalyptus* caused by *Elsinoë necatrix* sp. nov. Plant Pathol. 2021;70(5):1230–42.
6. Roux J, Wingfield MJ, Marincowitz S, Solís M, Phungula S, Pham NQ. Eucalyptus scab and shoot malformation: a new disease in South Africa caused by a novel species, *Elsinoë masingae*. Forestry: Int J For Res. 2023;97(2):327–38.
7. Swart L, Crous PW, Kang J-C, McHau GRA, Pascoe I, Palm ME. Differentiation of species of *Elsinoë* associated with scab disease of Proteaceae based on morphology, symptomatology, and ITS sequence phylogeny. Mycologia. 2001;93(2):366–79.
8. Weiss U, Ziffer H, Batterham TJ, Blumer M, Hackeng WH, Copier H, Salemink CA. Pigments of *Elsinoë* species. i. Pigment production by *Elsinoë* species; isolation of pure elsinochromes a, b, and c. Can J Microbiol. 1965;11:57–66.
9. Liao HL, Chung KR. Cellular toxicity of elsinochrome phytotoxins produced by the pathogenic fungus, *Elsinoë fawcettii* causing citrus scab. New Phytol. 2008;177(1):239–50.
10. Liao HL, Chung KR. Genetic dissection defines the roles of elsinochrome Phytotoxin for fungal pathogenesis and conidiation of the citrus pathogen *Elsinoë fawcettii*. Mol Plant Microbe Interact. 2008;21(4):469–79.
11. Jiao W, Liu L, Zhou R, Xu M, Xiao D, Xue C. Elsinochrome phytotoxin production and pathogenicity of *Elsinoë arachidis* isolates in China. PLoS ONE. 2019;14(6):e0218391.
12. Daub ME, Herrero S, Chung KR. Photoactivated perylenequinone toxins in fungal pathogenesis of plants. FEMS Microbiol Lett. 2005;252(2):197–206.
13. Lund NA, Robertson A, Whalley WB. The chemistry of fungi. Part XXI. Asperxanthone and a preliminary examination of aspergillin. J Chem Soc. 1953;494:2434–9.
14. Liu WZ, Shen YX, Liu XF, Chen YT, Xie JL. A new perylenequinone from *Hypomyces* sp. Chin Chem Lett. 2001;12:431–2.
15. Cordero RJ, Casadevall A. Functions of fungal melanin beyond virulence. Fungal Biol Rev. 2017;31(2):99–112.
16. Belozerskaya TA, Gessler NN, Aver'yanov AA. Melanin pigments of fungi. In: Merillon JM, Ramawat KG, editors. Fungal Metabolites. Cham: Springer International Publishing; 2016. p. 1–29.
17. Suthar M, Dufossé L, Singh SK. The enigmatic world of fungal melanin: a comprehensive review. J Fungi (Basel). 2023;9(9):891.
18. Hu J, Sarrami F, Li H, Zhang G, Stubbs KA, Lacey E, Stewart SG, Karton A, Piggott AM, Chooi YH. Heterologous biosynthesis of elsinochrome A sheds light on the formation of the photosensitive perylenequinone system. Chem Sci. 2019;10(5):1457–65.
19. Sabatini M, Comba S, Altabe S, Recio-Balsells AI, Labadie GR, Takano E, Gramajo H, Arabolaza A. Biochemical characterization of the minimal domains of an iterative eukaryotic polyketide synthase. Febs j. 2018;285(23):4494–511.
20. Chung KR, Liao HL. Determination of a transcriptional regulator-like gene involved in biosynthesis of elsinochrome phytotoxin by the citrus scab fungus, *Elsinoë fawcettii*. Microbiol (Reading). 2008;154(Pt 11):3556–66.
21. Ebert MK, Spanner RE, de Jonge R, Smith DJ, Holthusen J, Secor GA, Thomma B, Bolton MD. Gene cluster conservation identifies melanin and perylenequinone biosynthesis pathways in multiple plant pathogenic fungi. Environ Microbiol. 2019;21(3):913–27.
22. Jeffress S, Arun-Chinnappa K, Stodart B, Vaghefi N, Tan YP, Ash G. Genome mining of the citrus pathogen *Elsinoë fawcettii*; prediction and prioritisation of candidate effectors, cell wall degrading enzymes and secondary metabolite gene clusters. PLoS ONE. 2020;15(5):e0227396.
23. Jiao W, Xu M, Zhou R, Fu Y, Li Z, Xue C. Genomic analysis of *Elsinoë arachidis* reveals its potential pathogenic mechanism and the biosynthesis pathway of elsinochrome toxin. PLoS ONE. 2021;16(12):e0261487.
24. Wingfield BD, Berger DK, Coetzee MPA, Duong TA, Martin A, Pham NQ, van den Berg N, Wilken PM, Arun-Chinnappa KS, Barnes I, et al. IMA genome-F17. IMA Fungus. 2022;13(1):19.
25. Pham NQ, Duong TA, Wingfield BD, Barnes I, Durán A, Wingfield MJ. Characterisation of the mating-type loci in species of *Elsinoë* causing scab diseases. Fungal Biol. 2023;127(12):1484–90.
26. Cequeña RCM, Sumabat-Dacones LG. Presence of elsinochrome and other putative effectors in select genomes of the plant pathogen *Elsinoë* spp. based on in silico analysis. SciEnggJ. 2024;17(1):59–70.
27. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, Connor R, Funk K, Kelly C, Kim S, et al. Database resources of the national center for biotechnology information. Nucleic Acids Res. 2022;50(D1):D20–d26.
28. The Galaxy platform for accessible. reproducible, and collaborative data analyses: 2024 update. Nucleic Acids Res. 2024;52(W1):W83–w94.
29. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29(8):1072–5.
30. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31(19):3210–2.
31. Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome Res. 2008;18(12):1979–90.
32. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28(12):1647–9.
33. Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, Weber T. antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res. 2021;49(W1):W29–w35.
34. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403–10.

35. Li Z, Fan Y, Chang P, Gao L, Wang X. Genome sequence resource for *Elsinoë ampelina*, the causal organism of grapevine anthracnose. Mol Plant-Microbe Interact. 2020;33(4):576–9.
36. Zhang X, Zou H, Yang Y, Fang B, Huang L. Genome Resource for *Elsinoë batatas*, the causal agent of stem and foliage scab disease of sweet potato. Phytopathology. 2022;112(4):973–5.
37. Shanmugam G, Jeon J, Hyun J-W. Draft Genome sequences of *Elsinoë fawcettii* and *Elsinoë australis* causing scab diseases on citrus. Mol Plant-Microbe Interact. 2020;33(2):135–7.
38. Haridas S, Albert R, Binder M, Bloem J, LaButti K, Salamov A, Andreopoulos B, Baker SE, Barry K, Bills G, et al. 101 Dothideomycetes genomes: a test case for predicting lifestyles and emergence of pathogens. Stud Mycol. 2020;96:141–53.
39. Gostinčar C, Ohm RA, Kogej T, Sonjak S, Turk M, Zajc J, Zalar P, Grube M, Sun H, Han J, et al. Genome sequencing of four *Aureobasidium pullulans* varieties: biotechnological potential, stress tolerance, and description of new species. BMC Genomics. 2014;15:549.
40. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30(9):1236–40.
41. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30(14):3059–66.
42. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33(7):1870–4.
43. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30(9):1312–3.
44. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. New Orleans; 2010. p. 1–8. https://doi.org/10.1109/GCE.2010.5676129.
45. Gilchrist C, Chooi YH. clinker & clustermap.js: Automatic generation of gene cluster comparison figures. Bioinformatics. 2020;37(16):2473–5 https://cagecat.bioinformatics.nl/tools/clinker.
46. Collemare J, Billard A, Böhnert HU, Lebrun MH. Biosynthesis of secondary metabolites in the rice blast fungus *Magnaporthe grisea*: the role of hybrid PKS-NRPS in pathogenicity. Mycol Res. 2008;112(Pt 2):207–15.
47. Ren J, Wen L, Gao X, Jin C, Xue Y, Yao X. DOG 1.0: illustrator of protein domain structures. Cell Res. 2009;19(2):271–3.
48. Christiansen JV, Isbrandt T, Petersen C, Sondergaard TE, Nielsen MR, Pedersen TB, Sørensen JL, Larsen TO, Frisvad JC. Fungal quinones: diversity, producers, and applications of quinones from *Aspergillus*, *Penicillium*, *Talaromyces*, *Fusarium*, and *Arthrinium*. Appl Microbiol Biotechnol. 2021;105(21):8157–93.
49. Fan Y, Liu X, Keyhani NO, Tang G, Pei Y, Zhang W, Tong S. Regulatory cascade and biological activity of *Beauveria bassiana* oosporein that limits bacterial growth after host death. Proc Natl Acad Sci U S A. 2017;114(9):E1578–e1586.
50. Uchimiya M, Stone AT. Reversible redox chemistry of quinones: Impact on biogeochemical cycles. Chemosphere. 2009;77(4):451–8.
51. Daub ME, Herrero S, Chung KR. Reactive oxygen species in plant pathogenesis: the role of perylenequinone photosensitizers. Antioxid Redox Signal. 2013;19(9):970–89.
52. Mazur M, Masłowiec D. Antimicrobial activity of lactones. Antibiotics (Basel). 2022;11(10):1327.
53. Evidente A. Fungal bioactive macrolides. Nat Prod Rep. 2022;39(8):1591–621.
54. Lenz KD, Klosterman KE, Mukundan H, Kubicek-Sutherland JZ. Macrolides: From toxins to therapeutics. Toxins (Basel). 2021;13(5):347.
55. Markham JE, Hille J. Host-selective toxins as agents of cell death in plant–fungus interactions. Mol Plant Pathol. 2001;2(4):229–39.
56. Pusztahelyi T, Holb IJ, Pócsi I. Secondary metabolites in fungus-plant interactions. Front Plant Sci. 2015;6:573.
57. Weiss U, Flon H, Burger WC. The photodynamic pigment of some species of *Elsinoë* and *Sphaceloma*. Arch Biochem Biophys. 1957;69:311–9.
58. Wisecaver JH, Rokas A. Fungal metabolic gene clusters-caravans traveling across genomes and environments. Front Microbiol. 2015;6:161.
59. Hurst LD, Pál C, Lercher MJ. The evolutionary dynamics of eukaryotic gene order. Nat Rev Genet. 2004;5(4):299–310.
60. Keller NP, Turner G, Bennett JW. Fungal secondary metabolism – from biochemistry to genomics. Nat Rev Microbiol. 2005;3(12):937–47.
61. Wong S, Wolfe KH. Birth of a metabolic gene cluster in yeast by adaptive gene relocation. Nat Genet. 2005;37(7):777–82.
62. Cary JW, Ehrlich KC. Aflatoxigenicity in *Aspergillus*: molecular genetics, phylogenetic relationships and evolutionary implications. Mycopathologia. 2006;162(3):167–77.
63. Batada NN, Urrutia AO, Hurst LD. Chromatin remodelling is a major source of coexpression of linked genes in yeast. Trends Genet. 2007;23(10):480–4.
64. Khaldi N, Collemare J, Lebrun MH, Wolfe KH. Evidence for horizontal transfer of a secondary metabolite gene cluster between fungi. Genome Biol. 2008;9(1):R18.
65. Sieber CM, Lee W, Wong P, Münsterkötter M, Mewes HW, Schmeitzl C, Varga E, Berthiller F, Adam G, Güldener U. The *Fusarium graminearum* genome reveals more secondary metabolite gene clusters and hints of horizontal gene transfer. PLoS ONE. 2014;9(10):e110311.
66. Mann CWG, Sawyer A, Gardiner DM, Mitter N, Carroll BJ, Eamens AL. RNA-based control of fungal pathogens in plants. Int J Mol Sci. 2023;24(15):12391.

Publisher's Note