

RESEARCH

Open Access



# Genome analysis of *Cephalotrichum gorgonifer* and identification of the biosynthetic pathway for rasfonin, an inhibitor of KRAS dependent cancer

Andreas Schüller<sup>1</sup>, Lena Studt-Reinhold<sup>1</sup>, Harald Berger<sup>1</sup>, Lucia Silvestrini<sup>1,9</sup>, Roman Labuda<sup>2,3</sup>, Ulrich Güldener<sup>4,10</sup>, Markus Gorfer<sup>5</sup>, Markus Bacher<sup>2,6</sup>, Maria Doppler<sup>7,8</sup>, Erika Gasparotto<sup>1,2,11</sup>, Arianna Gattesco<sup>1,2</sup>, Michael Sulyok<sup>7</sup> and Joseph Strauss<sup>1,2\*</sup>

## Abstract

**Background** Fungi are important sources for bioactive compounds that find their applications in many important sectors like in the pharma-, food- or agricultural industries. In an environmental monitoring project for fungi involved in soil nitrogen cycling we also isolated *Cephalotrichum gorgonifer* (strain NG\_p51). In the course of strain characterisation work we found that this strain is able to naturally produce high amounts of rasfonin, a polyketide inducing autophagy, apoptosis, necroptosis in human cell lines and showing anti-tumor activity in KRAS-dependent cancer cells.

**Results** In order to elucidate the biosynthetic pathway of rasfonin, the strain was genome sequenced, annotated, submitted to transcriptome analysis and genetic transformation was established. Biosynthetic gene cluster (BGC) prediction revealed the existence of 22 BGCs of which the majority was not expressed under our experimental conditions. In silico prediction revealed two BGCs with a suite of enzymes possibly involved in rasfonin biosynthesis. Experimental verification by gene-knock out of the key enzyme genes showed that one of the predicted BGCs is indeed responsible for rasfonin biosynthesis.

**Conclusions** This study identified a biosynthetic gene cluster containing a key-gene responsible for rasfonin production. Additionally, molecular tools were established for the non-model fungus *Cephalotrichum gorgonifer* which allows strain engineering and heterologous expression of the BGC for high rasfonin producing strains and the biosynthesis of rasfonin derivatives for diverse applications.

**Keywords** Secondary metabolism, *Cephalotrichum*, *Doratomyces NG\_p51*, Rasfonin, Transcriptome analysis, Genome analysis, In silico retrosynthesis, Biosynthetic gene cluster prediction, Natural products discovery

\*Correspondence:

Joseph Strauss

[joseph.strauss@boku.ac.at](mailto:joseph.strauss@boku.ac.at)

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



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

### Fungi as producers of bioactive substances

Filamentous fungi have the potential to produce a countless number and huge diversity of bioactive substances, also called secondary metabolites (SMs). Those substances fulfil many different functions in the ecology and lifecycles of fungi. They are produced in response to certain developmental stages, metabolic or environmental conditions and stresses, like extreme temperatures [1], osmotic stress [2], UV-radiation [3], nutrient starvation [4, 5], in defence to other organisms such as competitors [6], mating partners [7] and during host–pathogen interactions as signalling molecules or virulence factors [8]. SMs are thus molecules that are not essential for viability, but provide a selective advantage in specific situations during the life cycle of a fungal cell or colony. This is also the reason why almost all of these substances are only produced by fungi when a specific developmental, metabolic or environmental signal is present and genetically received. While some SMs can be potentially used as pharmaceuticals (i.e., lovastatin (e.g., *Pleurotus ostreatus*) [9], penicillin (*Penicillium* sp.) [10]), others can have detrimental effects as toxic food contaminant (i.e., aflatoxins (*Aspergillus flavus*) [11], fumonisins or deoxynivalenol (*Fusarium* sp.) [12, 13], etc.). Therefore, SMs are of major ecological, economical and medicinal importance.

Bioinformatic analysis suggests that fungal genomes typically hold the genetic information for dozens of so-called biosynthetic gene clusters (BGCs) that serve as blue prints for single SMs [14], and even more than 90 different BGCs in a single species have been reported [15]. This genetic diversity gives rise to an enormous variety of SMs, even not considering the numerous intermediate products that are being produced on the way [16]. This makes fungi sheer endless reservoirs for novel bioactive substances, given that there are estimated to be more than 5 million fungal species worldwide [17].

During a project that focused on the role of fungi within the nitrate assimilation process in agricultural soil, a community analysis was performed that elucidated the composition of species and the phylogenetic relationship of the genes responsible for nitrogen assimilation. For a few isolates, among them *Cephalotrichum gorgonifer* NG\_p51 (previously classified as *Doratomyces* sp.), an unusual placement of the nitrate assimilation genes has been observed [18, 19]. This led to the conclusion that there had to be an, until then, unknown way of acquisition of the complete nitrate assimilation cluster including the nitrate transporter, the nitrate reductase and the nitrite reductase by horizontal gene transfer. It was proposed that the cluster was transferred from a donor at the base of the Pezizomycotina to a recipient in the Microascales, to which *C. gorgonifer* belongs [20]. Clustering of

the three genes was retained, which is otherwise unusual in the Sordariomycetes. These findings prompted us to investigate also possible acquisitions of BGCs coding for SMs from other fungi to *C. gorgonifer*.

Since its first discovery in *Aspergillus nidulans* [21] it is now well established that the reason for no expression of BGCs under standard laboratory conditions is chromatin-based silencing [21–25]. We and others have thus used genetic or chemical chromatin modification to leverage the production of novel SMs in a number of fungi (e.g., treatment of *Doratomyces* sp. with the HDAC inhibitor valproic acid [26], deletion of a chromatin remodeler-encoding gene, *kdmB* in *Aspergillus nidulans* [27, 28], removal of heterochromatic marks in *Fusarium fujikuroi* [29], etc.). For recent reviews on methods to active silence BGCs for SMs see [25] and [30]. Using a medium throughput effort, we have tested hundreds of isolates from different environmental screening campaigns for their capacity to produce novel SMs upon chromatin activation. In the case of *C. gorgonifer* NG\_p51, this led to the observation of bacterial growth inhibition and the discovery of several additionally produced compounds [26].

Further investigation in our in-house metabolite screening facility [31] revealed that *C. gorgonifer* is able to produce rasfonin, an SM initially isolated from *Talaromyces* strains and found to induce apoptosis in Ras-dependent tumor cells [32–35]. Rasfonin carries the potential of an anti-cancer pharmaceutical as it is active against several tumour types. It was shown to suppress proliferation of pancreatic cancer cells with mutated KRas and also reduced clone formation, invasion and migration of cells with mutated K-Ras. Although the exact mode of action is unknown, rasfonin was found to reduce the expression of Son of sevenless (Sos1), a nucleotide exchange factor which is involved in RAS activation [34]. In renal cancer cells, rasfonin promotes cell apoptosis and autophagy in connection with the protein Akt (protein Kinase B) which is a frequently activated oncoprotein in tumour cells [33]. Also, other studies found a connection between apoptosis and rasfonin treatment e.g., in osteosarcoma [35], renal adenocarcinoma cell lines [32] and in rasfonin-induced necroptosis [32].

Although rasfonin has the potential of being a future pharmaceutical, the biosynthetic pathway is still not known. This, however, would greatly assist further research on this SM and facilitate the generation of rasfonin over-production strains for large scale applications. We propose here the biosynthetic pathway for rasfonin in *C. gorgonifer* NG\_p51 based on the presence and predicted functions of *rsf* cluster genes. Furthermore, we established state-of-the-art molecular biology methods and performed various “omics” analysis for further

strain and species characterisation i.e., genome analysis, BGC prediction, transcriptome analysis, morphological description, genetic transformation and CRISPR/Cas-mediated genome editing.

## Results

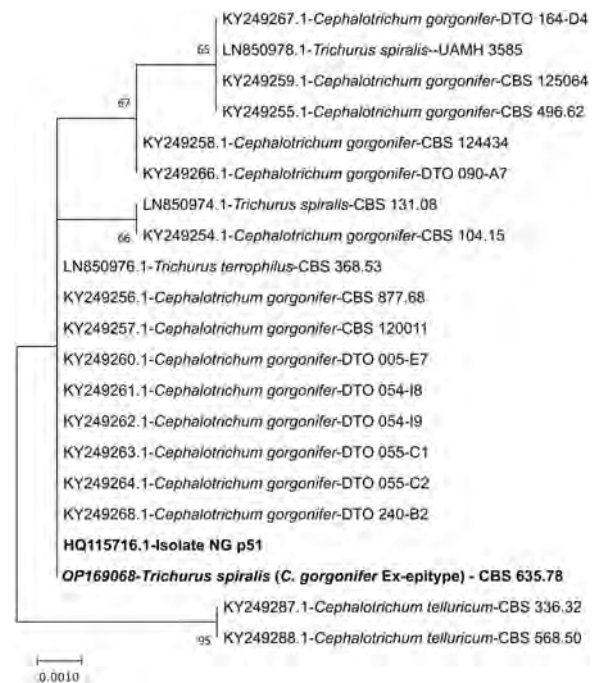
### Strain characterisation of *Cephalotrichum gorgonifer* NG\_p51

#### Ecology and taxonomy

We previously isolated the fungal strain NG\_p51 and according to the former taxonomic positioning determined the genus of the strain by its ITS region as *Doratomyces* sp. (GenBank: HQ115716.1) [18]. In the meantime, the genus *Cephalotrichum* (Microascaceae, Hypocreales) has been reorganized and now contains also species formerly affiliated to the genera *Doratomyces* and *Trichurus*. No sexual morphs have been observed so far. *C. gorgonifer* is frequently found in soil, decaying plant materials, dung and on wet cellulosic materials indoors. Occasionally it has also been isolated from human hair and respiratory samples. Their temperature optimum is around 25–30 °C but also growth at 37 °C has been observed for several isolates. There are, however, no indications for human pathogenicity but data in this regard is scarce [36, 37]. We performed a phylogenetic analysis with ITS sequences (=Internal transcribed spacer regions of the rDNA and 5.8S region) of strain NG\_p51 (GenBank No. HQ115716.1), together with the *C. gorgonifer* (including the Ex-epitype of *C. gorgonifer*: *Trichurus spiralis* CBS 635.-78) and *Cephalotrichum telluricum* (i.e.: closest out-group) dataset as published by Woudenberg et al. [37]. The resulting phylogenetic tree shows that our strain clearly belongs to the species *C. gorgonifer* and that the *C. telluricum* species are clustered as an outgroup as expected (Fig. 1). We therefore designated the strain *C. gorgonifer* NG\_p51.

#### Morphology of *C. gorgonifer* NG\_p51

Colonies of *C. gorgonifer* NG\_p51 reach a diameter between 35 and 55 mm on oat meal agar (OA), potato carrot agar (PCA) and potato dextrose agar (PDA) after 14 days at 25 °C. They effuse a dark greyish black (Fig. 2a) mycelium, consisting of nearly velutinous layer of synnemata which are 1–2 mm in length (Fig. 2b) and formed by aggregation of conidiophores (the annelophores). Numerous curled sterile setae are present alongside the fertile part of the synnemata (Fig. 2c, d) and conidia are borne from ampuliform annelids which appear smooth-walled and ovoidal to broadly ellipsoidal with truncate base and rounded apex, measuring 4–6×3.5–4.0 µm (Fig. 2e). Phenotypic traits of the strain NG\_p51 (especially presence of undulate sterile setae, greyish color of conidia *en-masse* as well as their shape and dimensions)

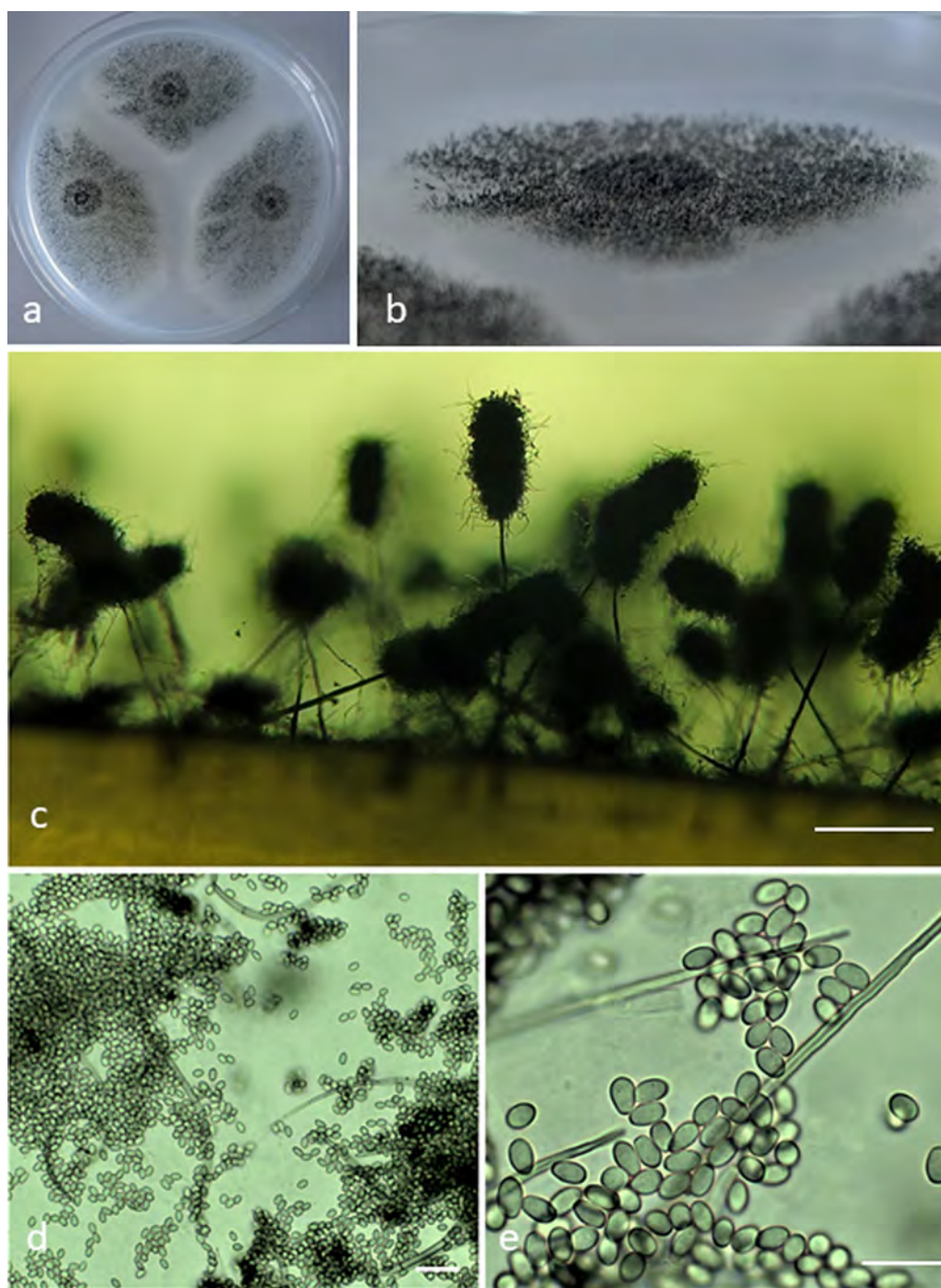


**Fig. 1** Phylogenetic tree of *Cephalotrichum gorgonifer* isolate NG\_p51. **A** Phylogenetic tree based on the ITS sequences of strain isolate NG\_p51 together with several *Cephalotrichum gorgonifer* isolates and *C. telluricum* strains as closest. Tree entry structure: "accession number—genus and species—strain identifier". The phylogenetic analysis shows that strain isolate NG\_p51 belongs to the species *Cephalotrichum gorgonifer*. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The strain isolate (i.e. NG\_p51) used for all experiments in this publication and the Ex-epitype are written in bold letters

are in perfect concordance with the species concept of *C. gorgonifer* (Bainier) Sandoval-Denis, Gené and Guarro [36], in older literature also described under a name *Trichurus spiralis* Hasselbr. by Domsch et al. [38] and Ellis [39]. Additionally, the growth on malt extract agar plates (MEA; n=3) at 28 °C, 30 °C, 32.5 °C, 35 °C and 37 °C showed that NG\_p51 behaves as published for *C. gorgonifer* [36] (Additional file 1: Fig. S10).

#### Secondary metabolite (SM) analysis of *C. gorgonifer*

Initial work with NG\_p51 revealed that it is capable of suppressing growth of *Staphylococcus aureus* and two clinical methicillin-resistant isolates of *S. aureus* when the fungal strain is grown in the presence of the HDAC inhibitor valproic acid. Valproic acid treatment increased the production of seven antimicrobial compounds (cyclo-(l-proline-L-methionine), p-hydroxybenzaldehyde, cyclo-(phenylalanine-proline), indole-3-carboxylic acid, phenylacetic acid and indole-3-acetic acid) and also led to the biosynthesis of an otherwise absent compound which was identified as phenyllactic acid [26]. Also, rasfonin has



**Fig. 2** Morphologies of *Cephalotrichum gorgonifer* NG\_p51. **a, b** Colonies on oatmeal agar (OA) after 14 d at 25 °C (Petri dish = 9 cm diam). **c** Synnemata in situ. **d, e** setae and conidia. Scale bars: **c** = 200  $\mu$ m, **d** = 20  $\mu$ m, **e** = 10  $\mu$ m

been found in these screens, although the production of this bioactive compound was subsequently shown not to be dependent on valproic acid in the growth medium [31]. Initial analysis of the production conditions revealed that this metabolite is produced on liquid *Aspergillus*

minimal medium (AMM) and detectable within 48 h of incubation (Fig. 5). Although rasfonin is a well-known bioactive compound with potential applications in cancer treatment [32–35], its biosynthetic pathway remains unknown. To decipher its production pathway, we

sequenced the genome and searched for the presence of BGCs potentially coding for rasfonin production.

### Genome analysis and gene annotation

Whole genome shotgun sequencing of *C. gorgonifer* was performed by 454 pyrosequencing yielding 1.78 Gb of raw sequence data which were assembled into 48 scaffolds of overall 35.9 Mb (49.4-fold coverage, N50 of 2.07 Mb). A total of 10,469 gene models were predicted using two different GeneMark gene predictors and manual validation. The annotated ORFs account for 43.8% of the genome. The overall GC content is 55%, while the average GC content of ORFs is 60.3%. Genome sequence data and annotation was deposited in the European Nucleotide Archive (ENA) at EMBL-EBI with the project number PRJEB15373 which can be accessed online at [40].

### Annotation of biosynthetic gene clusters (BGCs)

From one [41] to around 25 genes [42] can be involved in the biosynthesis of a single SM. In most cases these genes are coregulated and in physical vicinity to each other forming so-called BGCs [43]. Genes within these clusters can code for proteins that are responsible for the biosynthesis of the compound-defining backbone (e.g., polyketide synthase (PKS), non-ribosomal peptide synthetase (NRPS), terpene synthase (TPS)), the modification of the chemical backbone structure (e.g., cytochrome P 450, transferases, oxidoreductases, O-methyltransferases), the regulation of the cluster genes (global or cluster specific transcription factors (TFs)) or for the transport of synthesized metabolites (i.e., delivery to compartments within the cell or export to surrounding media) [14, 44].

For BGC annotation, antiSMASH 6.0.1 for fungi (fungiSMASH) was used with standard parameters [45]. The prediction algorithm for fungal BGCs assigned 308 genes to 24 putative BGCs. Six of them showed homology to BGCs from other species with similarities ranging from 20 to 100%. They are included in Table 1 with their respective references (only the best scoring homologous BGC is shown). Similarity in this case is defined as “percentage of genes within the closest known compound that have a significant BLAST hit to genes within the current region” (taken from antiSMASH documentation [45]). In total, ten PKSs and one type3-PKS, three NRPS-PKS hybrids, five NRPSs, six NRPS-like proteins and three putative TPSs could be assigned.

### Analysis of candidate genes by Whole genome and transcriptome sequencing under different physiological conditions

Because rasfonin is produced not during active growth but in stationary phase cultures, we compared expression

profiles between these two conditions. To obtain information about the genome-wide differential expression pattern of BGC genes we cultivated the fungus under conditions of active growth in liquid shake cultures for 24 h in AMM and compared the expression pattern with cells in stationary phase after 48 h of growth on the same medium that, at this stage, is already depleted for nutrients [5]. Of all 10,469 annotated genes, 717 genes that were predicted not to be part of a secondary metabolism pathway by antiSMASH were up-regulated at 48 h and 818 were down-regulated. From the 308 genes that are putatively involved in secondary metabolism, 39 were up-regulated at the 48-h timepoint and 40 were down-regulated. The remaining 231 genes stayed unaffected. 8855 genes were not differentially regulated between the two timepoints which are 85% of the annotated genes.

The whole transcript-dataset can be retrieved from the additional materials (Additional file 2). The whole transcriptome was uploaded at NCBI's Gene Expression Omnibus (GEO) [46] under the accession number GSE217303.

To estimate the transcriptional activity of BGCs under the chosen conditions of active growth versus stationary phase, we categorized the transcriptional status of the key-enzymes within a given BGC into three states (Table 1 Column Expression; Cat.). The expression of category 1 genes is not detectable or falls below the arbitrarily set threshold of biological relevance (see caption of Table 1) at both time-points (i.e., 14 of 29 genes). Category 2 genes (3 genes) are expressed only at one time point (i.e., either 24 h or 48 h) and the expression of category 3 genes (12 genes) is above the threshold at both 24 h and 48 h, although their expression levels may vary between the two time points (7 genes with at least a two-fold difference between the time points). The individual expression levels, differential regulation between 24 and 48 h and their respective transcriptional category of all predicted key-enzyme encoding genes can be retrieved from Table 1.

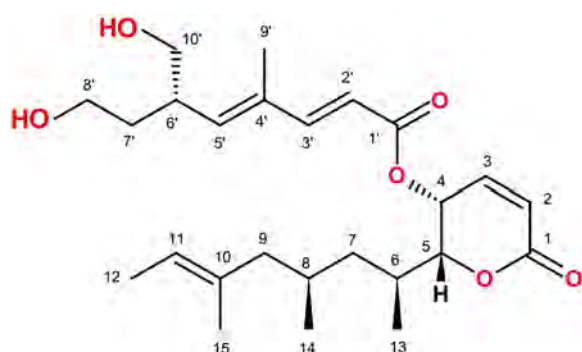
### Prediction of the BGC responsible for rasfonin biosynthesis based on PKS domain structure

Rasfonin (Fig. 3) is an  $\alpha$ -pyrone-containing natural product composed of two distinct acetate-based polyketide (PK) chains that are linked via an ester bond. The individual PK chains are likely to be derived from the condensation of four and six acetyl-CoA subunits resulting in the formation of a partially reduced and di- and trimethylated tetra- and hexaketide, respectively, exhibiting hydroxyl groups at C8' and C10'. These chemical features suggest the involvement of two highly-reducing PKS (HR-PKSs) containing, in addition to the essential acyltransferase (AT), ketoacyl-synthase (KS), and acyl

**Table 1** Summary of BGCs of *C. gorgonifer* NG\_p51 as predicted by antiSMASH

ID	BB-gene(s)	Type	Expression (BB- gene)				Homologs (Species)	Similarity	Ref.
			24h	48h	~	Cat.			
1	DNG_00901	terpene	5.6	6.2	0.6	Cat3	-	-	-
2	DNG_01262	T1PKS	5.3	5.6	0.3	Cat3	naphtalene BGC ( <i>Daldinia eschscholzii</i> )	33 %	(47)
3	DNG_01539	T1PKS	1.7	3.7	2.0*, +	Cat3	-	-	-
4	DNG_02059	T3PKS	4.9	6.1	1.2*	Cat3	-	-	-
5	DNG_02555	terpene	6.8	4.3	-2.5*, +	Cat3	-	-	-
6	DNG_02774	T1PKS	0.0	6.4	7.7*, +	Cat2	-	-	-
	DNG_02782	T1PKS	0.0	6.2	9.5*, +	Cat2		-	-
7	DNG_03072	T1PKS	0.0	0.0	-	Cat1	-	-	-
8	DNG_04293	NRPS	0.0	0.0	-	Cat1	-	-	-
	DNG_04303	T1PKS	0.0	0.0	-	Cat1	-	-	-
9	DNG_04582	NRPS-T1PKS	0.0	0.0	-	Cat1	betaenone A – C ( <i>Phoma betae</i> )	25 %	(48)
10	DNG_04746	terpene	7.7	7.5	-0.1	Cat3	Squalestatin ( <i>Aspergillus</i> sp.)	40 %	(49)
11	DNG_06431	T1PKS	0.0	0.0	-	Cat1	-	-	-
12	DNG_07105	T1PKS	0.0	0.0	-	Cat1	-	-	-
	DNG_07107	T1PKS	0.0	0.0	-	Cat1	-	-	-
13	DNG_07411	NRPS	2.1	2.7	0.6	Cat3	-	-	-
14	DNG_08044	NRPS	0.0	1.9	0.9	Cat2	phyllostictine A/B ( <i>Phyllosticta cirsii</i> )	20 %	(50)
	DNG_08052	NRPS-T1PKS	0.0	0.0	-	Cat1	-	-	-
15	DNG_08262	NRPS-like	0.0	0.0	-	Cat1	-	-	-
16	DNG_08309	NRPS	5.4	1.4	-4.1*, +	Cat3	Dimethylcoprogen ( <i>Alternaria alternata</i> )	100 %	(51)
17	DNG_09028	NRPS-like	5.2	5.2	0.0	Cat3	-	-	-
18	DNG_09136	NRPS-like	0.0	0.0	-	Cat1	-	-	-
19	DNG_09754	NRPS	0.0	0.0	-	Cat1	-	-	-
20	DNG_09868	NRPS-like	5.0	1.6	-3.4*, +	Cat3	-	-	-
21	DNG_09951	NRPS-like	0.0	0.0	-	Cat1	-	-	-
22	DNG_10196	NRPS-like	7.3	5.6	-1.7*	Cat3	-	-	-
23	DNG_10230	NRPS-T1PKS	0.0	0.0	-	Cat1	cytochalasin E / K ( <i>Aspergillus clavatus</i> )	23 %	(52)
24	DNG_10425	T1PKS	0.0	0.0	-	Cat1	-	-	-

Columns: ID: Cluster ID as assigned by antiSMASH; BB-gene(s): gene number of the key-enzyme encoding genes; type: type of key-enzyme; Expression: Expression of backbone genes at 24 h and 48 h are depicted as a 2-zone heatmap (weak expression | red → blue | strong expression), based on their RPKM value. An arbitrary log-value of 1 was considered as a biologically relevant expression threshold under which, the expression was set to zero. The column “Δ” depicts the differential expression between 24-h and 48-h samples and is visually supported by a 3-zone heatmap (down-regulated @ 48 h | blue → white → red | up-regulated @ 48 h) with zero being the turning point. A significance value of  $p < 0.05$  was applied as threshold which is indicated by an Asterisk near the differential expression value. If the differential expression is higher than twofold, it is indicated with an additional “+”. The Cat.: Expression Category (Cat1: Silent at 24 h and 48 h; Cat2: Expressed at either 24 h or 48 h (i.e., differentially expressed); Cat3 (Expressed at both 24 h and 48 h). Best scoring homolog BGCs in other species that were found by antiSMASH are also displayed together with their similarity in % and references



**Fig. 3** Structure of rasfonin

carrier protein (ACP) domains, domains involved in the reduction of the growing PK chain, i.e., a ketoreductase (KR), a dehydratase (DH) and an enoylreductase (ER). Next to this, the methyl groups at C6, C8, C10, and C4', C6', suggest the presence of an intrinsic C-methyltransferase (MT) domain that functions in the transfer of a methyl group from S-adenosylmethionine (SAM) to a  $\beta$ -ketoacyl-ACP substrate for both involved PKSs.

We thus screened the genome sequence of *C. gorgonifer* for the presence of predicted PKS-encoding genes that harbour a domain composition putatively fitting the enzymatic functions necessary for rasfonin biosynthesis. Bioinformatic analysis of the 14 PKSs present in the *C. gorgonifer* genome revealed that seven HR-PKSs harbour the necessary domain architecture (Fig. 4a). Only two BGCs contain two HR-PKSs in close proximity to each other namely BGC 6 (Table 1 ID 1) with the combination of [CgPKS4 + CgPKS5] (Fig. 4a) and BGC 12 (Table 1 ID 12) with the combination [CgPKS10 + CgPKS11] (Fig. 4a). Both also encode a putative transferase, i.e., DNG\_02776 and DNG\_07106, respectively. To further define which of the two BGCs may be involved in rasfonin biosynthesis, we questioned our transcriptome data set and found that only the BGC with the HR-PKS pair *CgPKS4* and *CgPKS5* (Fig. 4a; Table 1 ID 6) showed expression at 48 h, the stationary phase condition conducive for rasfonin production. The other candidate gene pair *CgPKS10* and *CgPKS11* (Fig. 4a; Table 1—ID 12) was not expressed at any tested condition.

Next, we set out to analyse the extension of the CgPKS4/CgPKS5 BGC. Our transcriptome data showed a clear coregulation in expression of the genes that are positioned between *CgPKS4* and *CgPKS5*. No transcription was observed during active growth (24 h post inoculation, hpi) while expression of all putative cluster genes was significantly upregulated 48 hpi. The genes upstream of CgPKS4 and downstream of CgPKS5 were not coregulated (Fig. 4b). The GO annotations of adjacent genes

do not suggest that they are involved in the synthesis of rasfonin. Thus, the BGC likely spans nine genes, i.e., DNG\_02774–DNG\_02782. To evaluate the contribution of the putative CgPKS4/CgPKS5 BGC to rasfonin biosynthesis, one of the PKS-encoding genes, *CgPKS4*, was chosen for targeted gene disruption.

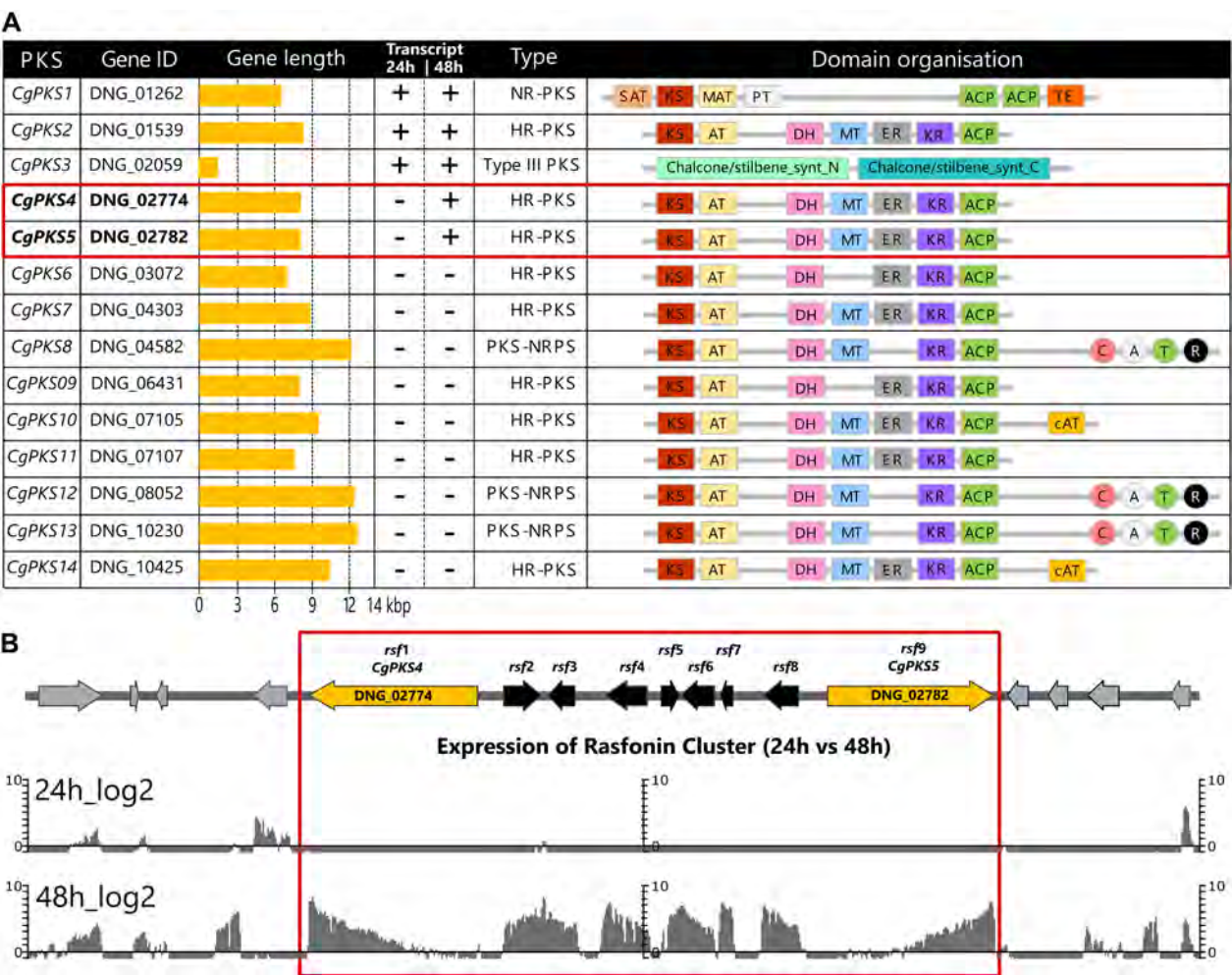
### Experimental verification of the BGC involved in rasfonin biosynthesis

The standard method to prove the involvement of a respective candidate gene in SM biosynthesis is deletion of the coding region in the genome and the analysis if the deletion genotype is accompanied by the loss of product formation. However, so far *Cephalotrichum* sp. have not yet been genetically modified and used for reverse genetics approaches. Therefore, a transformation method had to be established with the aim to disrupt or mutate *CgPKS4* and the test by chemical analysis if rasfonin is not produced any longer in the *CgPKS*-deleted strain. To maximize the chance obtaining *CgPKS4* mutations by transformation, three different approaches were selected, i.e., a Cas9 approach (Strain Cg-Cas9-02774-1), a transformation approach mediated by *Agrobacterium tumefaciens* (Strain Cg-At-02774-1) and also a homology directed repair (HDR) approach via protoplast-mediated transformation of a linear fragment (Strain Cg-HDR-02774-1) [53, 54]. All three transformation methods were successfully implemented in our *C. gorgonifer* strain NG\_p51. Cas9 and the HDR transformations were performed on cryo-stocks of *C. gorgonifer* protoplasts (see materials and methods section).

Disruption of *CgPKS4* by Cas9 resulted in the strain Cg-Cas9-02774-1. Sequencing of the locus revealed a 640 bp deletion ranging from base pair 269 (i.e.: sgRNA annealing site) to base pair 908 of the coding sequence (CDS), a mutation that additionally leads to a frameshift within the remaining CDS of the gene (Sequence at Cas9-restriction site: 5'-ACCAATGCACCCGGT | 640 bp deletion | GTTCGACCACCGCGC-3') (see Additional file 3).

The *Agrobacterium*-mediated mutation resulted in the transformant Cg-At-02774-1 that has the first 2300 base pairs of *CgPKS4* (DNG\_02774) replaced with the hygromycin resistance cassette (hph), again resulting in a frameshift within the CDS. The 5'-end homology region spans 1200 base-pairs upstream of the start codon (i.e. promoter region) and the 3'-homology region spans 1500 base-pairs downstream of the exchanged fragment (i.e. first 2300 base pairs of the gene body) (see Additional file 3).

The transformant Cg-HDR-02774-1 resulted from the HDR-directed approach. Here, the same replacement cassette was used as for the *Agrobacterium*-mediated

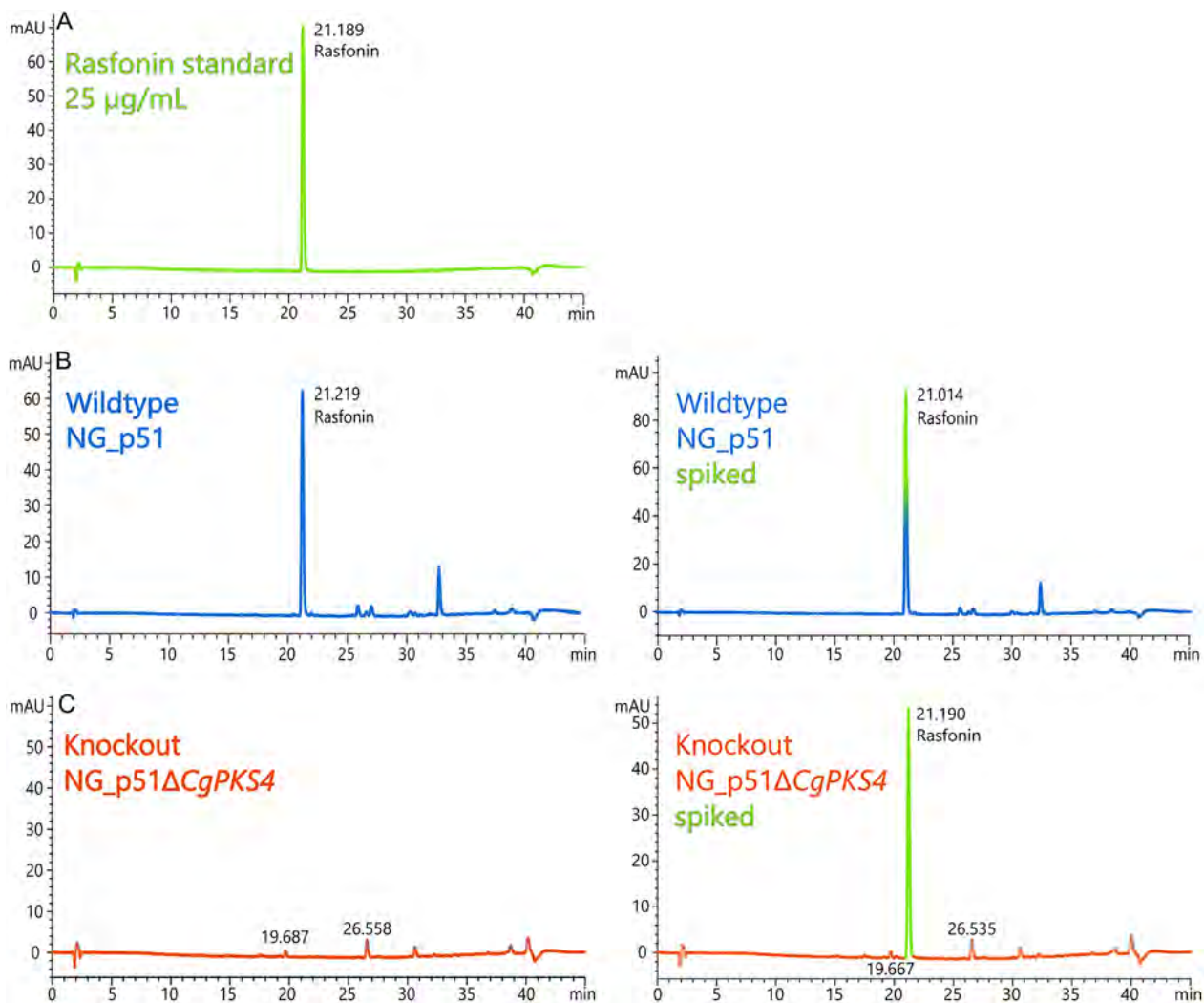


**Fig. 4** **A** Domain organization was analysed using the NCBI Conserved Domain [36], InterPro [37], SBSPKSV2 [38]; and the PKS/NRPS Analysis Web-site [39]. KS (red), keto synthase; AT (yellow), acyltransferase; DH (pink), dehydratase; MT (blue) C-methyltransferase; ER (grey) enoylreductase; KR (violet), ketoreductase; ACP (green), acyl carrier protein; CAT (orange), carnitine acyltransferase; Chalcone/stilbene\_synt\_N (lime-green): Chalcone/stilbene synthase N-terminal; Chalcone/stilbene\_synt\_C (dark-cyan): Chalcone/stilbene synthase C-terminal. The proposed PKS names are listed together with their gene number and length. The transcriptional activity at 24 h or 48 h ins indicated as “-” (No transcription at given timepoint) or “+” (transcription at given timepoint). The two PKS encoding genes that are proposed to be involved in the rasfonin synthesis (i.e. CgPKS4 and CgPKS5) are highlighted by a red box **B** The proposed BGC for rasfonin biosynthesis is depicted with transcript at 24 h and 48 h timepoint (shake flask culture, AMM). The BGC borders are depicted in red and were assigned based on the coregulation of genes. Proposed BGC gene names (rsf1 to rsf9) are depicted above gene illustration

transformation (i.e., PCR product from the “HDR-Fragment” primer pair with template pAt-DNG02774; Additional file 1: Table S1). Both strains [i.e., Cg-At-02774-1 and Cg-HDR-02774-1] were PCR tested for the successful gene replacement event (see Additional file 3: Fig. S1; Additional file 3)].

Phenotypical analysis of the various *rsf1* deletion strains on MEA agar plates did not reveal any obvious morphological or growth phenotypes (data not shown). To test if the mutated target gene is responsible for the biosynthesis of rasfonin, the transformants and the

recipient strain NG\_p51 were cultivated in parallel on MEA agar medium. The fungal colonies (mycelia + agar media) were extracted, and extracts were analysed by high performance liquid chromatography (HPLC) for the presence of rasfonin. None of the three tested mutants showed any detectable amounts of rasfonin (Fig. 5c, left) when compared to the wildtype NG\_p51 (Fig. 5b, left). A 25 µg/mL rasfonin solution produced in our laboratory and verified by NMR (see materials and methods) served as analytical standard (Fig. 5a). Next to this, we spiked the wildtype sample with our



**Fig. 5** HPLC analysis of rasfonin between a 25 µg/mL rasfonin standard (green, **A**), the wildtype isolate NG\_p51 (blue, **B** left) and the CgPKS4 knock-out strain NG\_p51ΔCgPKS4 (orange, **C** left). All 3 knock-out strains and all replicates showed the same absence of rasfonin, but only one is shown here. Spiked samples are shown on the right panel and are indicated by a green peak (NG\_p51, **B** right; NG\_p51ΔCgPKS4, **C** right). The retention times in minutes are shown next to the peaks

standard thereby unequivocally proving that the peak in the wildtype chromatogram indeed is rasfonin (Fig. 5b, c, right), and that rasfonin is missing in our *CgPKS4* mutants. To verify that not even trace amounts of rasfonin are produced, that are not detected by our HPLC method, we subjected an extract from the Cg-Cas9-02774-1 strain to liquid chromatography-high resolution mass spectrometry (LC-HRMS). The results of this analysis showed no detectable amounts of rasfonin in the mutant strains (Additional file 3: Fig. S2). This verifies that *CgPKS4* (from now on referred to as Rsf1) is indeed directly involved in the biosynthesis of rasfonin in *C. gorgonifer* NG\_p51.

Next to *CgPKS4* (*rsf1*), eight genes are coregulated under rasfonin-producing conditions including a second PKS-encoding gene *CgPKS5* located at the end of the putative BGC. Thus, we propose that the BGC involved in rasfonin biosynthesis stretches between *CgPKS4* and *CgPKS5*, and comprises overall nine genes, *rsf1* to *rsf9* (Fig. 4b). Next to the PKSs, the BGC encodes three putative cytochrome p450s (DNG\_02775: *rsf2*; DNG\_02779: *rsf6*; DNG\_02781: *rsf8*), a putative *O*-acyl-transferase (DNG\_02776: *rsf3*), a Major facilitator family1 (MFS1) transporter (DNG\_02777: *rsf4*), a putative *trans*-thioesterase harbouring a FSH domain and an alpha/beta hydrolase fold (DNG\_02778: *rsf5*) as well as one uncharacterised protein (DNG\_02780: *rsf7*).

## Discussion

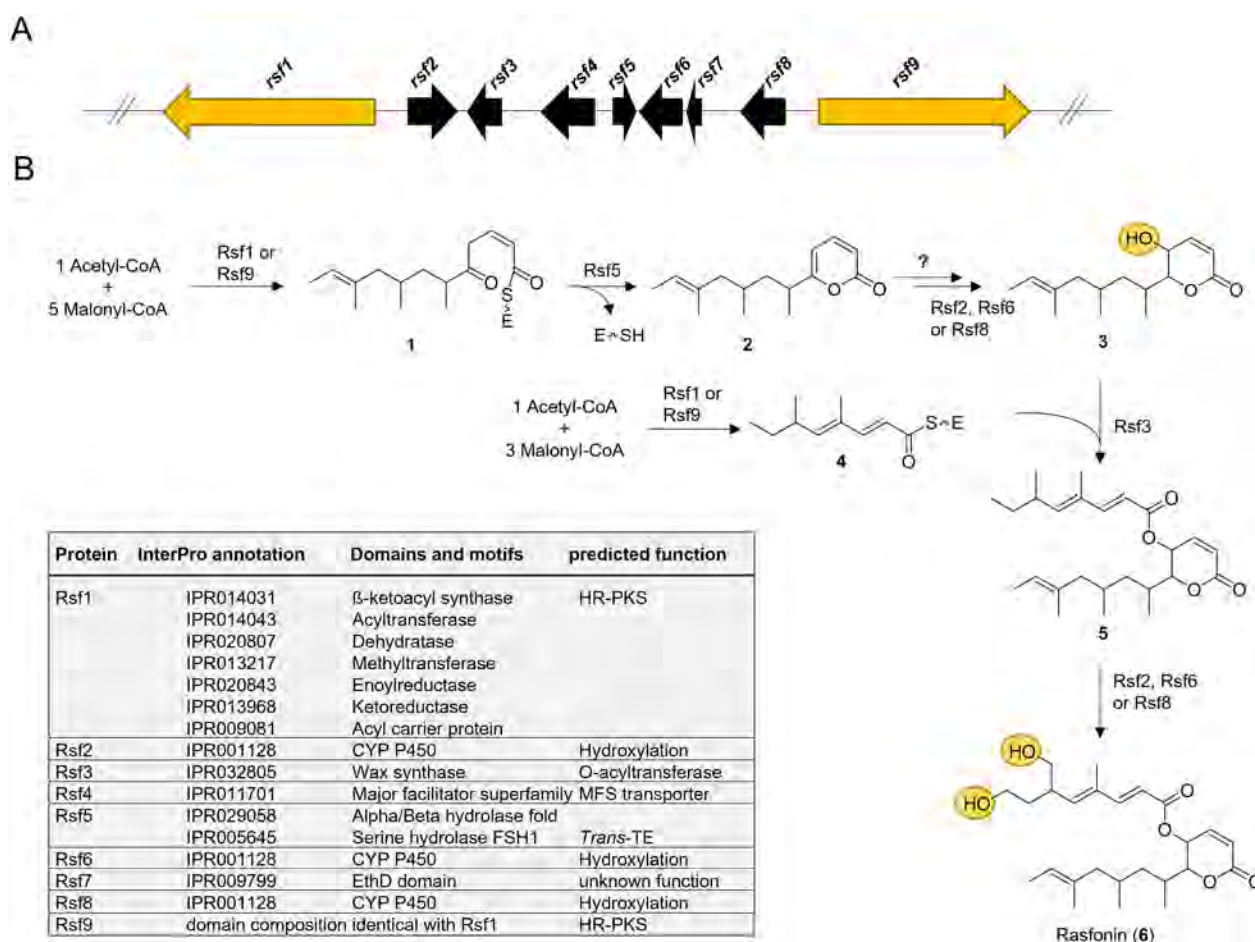
Using a combination of whole genome sequencing, bioinformatic analysis, genome-wide transcriptome analysis of cells producing versus non-producing conditions allowed us to narrow down the possible biosynthetic genes for rasfonin. These data formed the basis for prediction of enzymatic domains which are known to synthesize similar molecules. But all this prediction work in the end needs experimental verification by targeted inactivation of the proposed biosynthetic gene. For non-model fungi, the transformation-mediated inactivation of a target genes is often not trivial. Our strategy was to follow three well-known strategies to reach that goal with a high probability of success. But still, some modifications of standard protocols were necessary. Here, we developed a modified version of common protoplastation protocols to ensure a cost- and time-efficient method. The possibility to use cryo-stocks significantly increases the speed and flexibility of transformation experiments. As different fungi will need different conditions during protoplastation and post-transformation regeneration, we want to highlight a publication by Wu and co-workers [55] which compares the impact of different buffer compositions and conditions on protoplastation and regeneration efficiency. The publication assisted in the development of the protoplastation protocol for *C. gorgonifer* and will surely also be of use in case of establishment of other fungal species for the laboratory environment.

The biosynthesis of rasfonin is likely to be coordinated by two highly-reducing iterative type I PKSs. In general, two scenarios are possible that may yield rasfonin. In a first scenario one PKS generates a PK chain that serves as a starter unit and may be passed on to the second PKS for further extension. This mechanism has been proposed for zearalenone in *F. graminearum* [56], T-toxin in *Cochliobolus heterostrophus* [57], asperfuranone in *A. nidulans* [58] or botcinic acid in *Botrytis cinerea* [59]. In a second scenario two PKSs are involved in the assembly of two distinct PK chains that are subsequently linked via an ester bond catalysed by an acyltransferase encoded within the same BGC as it has been shown for lovastatin in *Aspergillus terreus* [60], and the related compactin in *Penicillium citrinum* [61]. Noteworthy, in rare cases one PKS is involved in the assembly of two distinct PK chains as it has been recently shown for gregatin A in *Penicillium* sp. Sh 18 [62], or fusamarin in *Fusarium mangiferae* [63], and which is likely also true for fujikurin in *Fusarium fujikuroi* [64, 65]. However, given the cluster architecture and co-regulation as well as the presence of two HR-PKS- and one putative transferase-encoding genes in the rasfonin BGC, DNG\_02776 (*rsf3*), we propose the following route for rasfonin biosynthesis (Fig. 6): One HR-PKS, *Rsf1* (CgPKS4) or *Rsf9* (CgPKS5), is involved

in the programmed assembly of the partially reduced 6,8,10-trimethylated hexaketide (**1**) by successive condensations of acetyl-CoA with 5 malonyl units, while the other assembles the 4',6'-dimethylated tetraketide (**4**). Noteworthy, none of the HR-PKSs harbours a release domain, but the presence of *Rsf5*, a predicted thioesterase (TE) with similarity to *LovG* from the lovastatin cluster [66], suggests that *Rsf5* may function in the release of the hexaketide. Similar *trans*-TEs have been shown to facilitate off-loading in the case of fusarielin biosynthesis in *F. graminearum* [67], and fusamarin biosynthesis in *F. mangiferae* [63]. The PK chain may be released either by TE-mediated hydrolysis followed by the spontaneous intramolecular cyclisation accompanied by dehydration yielding **2**, or the cyclisation happens directly through TE-catalysed pyrone formation as it has been shown for the TE-domain of CTB1 involved in cercosporin biosynthesis in *Cercospora* sp. [68]. **2** is then further modified to yield **3** by hydrogenation and subsequent hydroxylation of C4. The latter step is likely catalysed by one of the cytochrome p450s encoded within the BGC, *Rsf2*, *Rsf6* or *Rsf8*. Next, *Rsf3*, encoding a putative O-acyltransferase with similarity to *LovD* [69] captures **4**, followed by the transacylation of **3** resulting in the formation of **5**, that is further hydroxylated at C8' and C10' likely catalysed by one or both of the remaining cytochrome p450s encoded in the BGC, yielding the final product rasfonin (**6**).

In fungi, the true physiological or ecological function of rasfonin remains obscure. No strong antimicrobial activity has been detected so far. Available data from experiments with cancer cell lines suggest a modulation of Ras-signalling pathways by rasfonin, which can suppress proliferation of certain cell lines [34]. As a colonizer of cellulosic plant litter, *C. gorgonifer* would encounter numerous other eukaryotes, where alterations of their developmental program through modulation of Ras signalling could be potentially advantageous for the producing organism. However, other eukaryotic cells tested by us, such as *Saccharomyces cerevisiae*, seem highly resistant to this compound (our unpublished preliminary observations). The availability of KO strains without detectable rasfonin production will allow to test hypotheses on rasfonin function in the natural environment.

An interesting preliminary observation of our study was that the metabolite does not seem to be actively exported into the surrounding of the cell but stays within or attached to the mycelium in a fungal culture (unpublished preliminary observation). Although there is an expressed gene within the BGC that codes for a transporter, computational analysis of the predicted protein sequence by MULocDeep suggests that the transporter is targeted to a lysosome-like membrane [70]. This indicates that an intermediate or the final compound is stored



**Fig. 6** Proposed biosynthetic pathway of rasfonin. **A** The rasfonin BGC. **B** The proposed pathway of rasfonin synthesis

within extra compartments like it is the case with fungal toxismes [71, 72]. Of course, some exporter, encoded by a gene residing externally of the rasfonin BGC, might be responsible for export, but this topic needs further investigation.

## Conclusions

This work provided the tools and methods for working efficiently with the rasfonin producing strain *Cephalotrichum gorgonifer* NG\_p51 in a laboratory environment. Thorough analysis of the genome and different transcriptomes in combination with metabolic and taxonomic studies have allowed us to identify a BGC that is responsible for the biosynthesis of rasfonin. This now enables further train and pathway engineering with the aim to develop a high yield production strain of this pharmaceutically relevant compound. The establishment of three different transformation methods and the development of an adapted protocol for protoplast and protoplast-cryo-stock preparation of *C. gorgonifer* NG\_p51 surely helps to speed up molecular genetics work with this

fungus and will also help with establishing such methods in other fungal species promoting the identification of compounds encoded by so far orphan BGCs.

## Materials and methods

### Phylogenetic tree: evolutionary analysis by maximum likelihood method

ITS sequences (For accession numbers see Fig. 1) were aligned with the program MEGAX: Molecular Evolutionary Genetics Analysis across computing platforms (Version 10.2.6) [73] with the ClustalW algorithm (standard parameters) [74]. After alignment, sequences were trimmed to equal length for the construction of the phylogenetic tree.

The tree with the highest log likelihood (−808.77) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Tamura-Nei model [75], and then selecting the

topology with superior log likelihood value. This analysis involved 21 nucleotide sequences. There was a total of 557 positions in the final dataset. Evolutionary analyses were also conducted in MEGA X. Bootstrapping was performed 500 times.

#### Morphological analysis of strain NG\_p51

For phenotypic characterization, the strain *Cephalotrichum gorgonifer* NG\_p51 was transferred on Potato Dextrose Agar (PDA, Fluka), Malt Extract Agar (MEA, Merck) and Oatmeal Agar (OA) as described by [76] and incubated for 14 days in the dark at 25 °C. For comparative description of the macroscopic and microscopic characteristics, OA was used according to [36], and additionally compared to [38, 39] where the fungus is described and illustrated under the name *Trichurus spiralis* (current name: *Cephalotrichum gorgonifer* [36]). The photomicrographs were taken using a Motic BA 310 microscope with Motic Image Plus 3.0 software. Lactophenol blue was used as a mounting medium for microphotography. Photographs of the colonies were taken with a Sony DSC-RX100.

#### Growth phenotype at different temperatures

For determining the growth morphology under different temperatures, spores of strain NG\_p51 were grown by plating them from cyro-stocks (50% Tween-20 0.1% spore suspension and 50% glycerol 50% (w/v)) on MEA plates at 30 °C for 4 days. Spores were harvested by scraping off the agar with a spatula and 0.1% Tween solution. The spore suspension was then filtered through glass wool filled tips and a 1/1000 dilution was prepared (3 subsequent 1/100 dilutions; 100 µL spore solution and 900 µL Tween 0.1%) were then counted in a counting chamber (Neubauer improved; PC72.1—Carl Roth) and a dilution with a spore density of 100 spores/µL (Tween-20 0.1%) was prepared. 10 µL of the spore solution was then applied in the middle of a MEA plate and 3 plates each were incubated at 28 °C/30 °C/32.5 °C/35 °C/37 °C for 4 days. The growth was measured in mm diameter.

#### Genome sequencing and gene prediction

The genome sequencing and assembly of *Cephalotrichum gorgonifer* was performed using a whole-genome shotgun approach that explored paired-end 454 (LGC Genomics GmbH, Berlin, Germany). Using Newbler v2.6. (Roche (2010) 454 Sequencing System Software Manual, v 2.5.3: Part C—GS De Novo Assembler, GS Reference Mapper, SFF Tools, 454. Life Sciences Corp, A Roche Company, Branford) shotgun and 8 kb-PE library reads were assembled into 702 contigs. Pre-assembled contigs were combined into 48 scaffolds using the SSPACE algorithm [77]. To predict genes on the scaffolds two ab initio gene

predictor algorithms were applied, GeneMark-S and GenMark-ES version 2.3 [78]. Gene models differently predicted by the algorithms were manually curated. The genome was analysed using the PEDANT system [79] to allow comparative feature analysis.

#### Transcriptome: RNA sequencing and analysis

Total RNA extraction samples (24 h/48 h; 6 µg each) were transferred to Vienna Biocenter Core Facilities [80] for library preparation and Illumina high throughput sequencing using poly-A enrichment kit (NEB) and Nextera Library prep kit. 50 bp single end sequencing was performed using a HiSeq v4 Illumina sequencer. Obtained sequences were de-multiplexed, quality controlled, filtered using trimmomatic 0.36 [81] and mapped on the *Cephalotrichum gorgonifer* NG\_p51 genome assembly (ENA project number PRJEB15373). Mapping was performed using BWA [82]. Mapping data were processed using samtools [83] and duplicate reads were removed using Picard MarkDuplicates accessible at [84]. Reverse transcripts were counted using python script HTSeq [85]. Normalization and statistics were done using R/Bioconductor and the limma and edgeR packages, using mean–variance weighting (voom) and TMM normalisation [86] and robust linear model fitting. A significance cut-off of  $p < 0.05$  and an absolute difference of 1 (i.e. twofold regulation) was applied for analysis. Transcription levels are log2 read counts per kilobase of exon per million library reads (RPKM). For trace graphs transcript coverage was calculated using bedtools genomecov [87, 88] for each bp normalized to sequencing library sizes as log2 scaled counts per million reads (CPM). All data are available at NCBI GEO under the accession number GSE217303.

#### Molecular biology methods

For gDNA extraction, *Cephalotrichum gorgonifer* colonies were inoculated in 1 mL liquid Czapek DOX media with 10 mM NO<sub>3</sub> in Lysing Matrix A tubes (MP Biomedicals; Irvine, CA; Art. No. SKU 116910100) for 24–48 h at 30 degrees and 180 rpm. After incubation, the cultures were centrifuged at 20,238 rcf in an Eppendorf 5424 centrifuge. The supernatant was removed and 1 mL CENIS lysis buffer (200 mM TRIS–HCl pH 8.5, 250 mM NaCl, 25 mM EDTA, 0.5% SDS) was added. A Fast-Prep-24™ ribolyzer (MP Biomedicals; Irvine, CA; Art. No. SKU116004500) was then used to homogenize the samples (20 s, 6.0 m/s) and the gDNA extraction according to CENIS was conducted [89]. The dried gDNA was dissolved in 100 µL of dH<sub>2</sub>O and heated up to 65 °C for 10 min. 1 µL was taken for PCR.

Polymerase Chain Reaction (PCR) was either performed with the Q5®-polymerase from New England

Biolabs (New England Biolabs; Art. No. M0491L) for plasmid cloning, or with the GoTaq® Green Master Mix (Promega; Madison, WI, USA; Art. No. M7845) if used for diagnostic PCRs (i.e. strain and plasmid verification). PCR was performed according to the respective manual. Plasmid preparations were performed with GeneJet® Plasmid Miniprep Kit (Thermo Scientific™, Art. No. K0503). All information (primer pair names and sequences, amplicon names and templates) can be retrieved from Additional file 3: Table S1.

Plasmid construction was performed with either NEBuilder® HiFi DNA Assembly Master Mix (New England Biolabs (NEB); Art. No. E2621S) or by yeast recombinational cloning (YRC) [90]. All PCR fragments were purified with the Monarch PCR & DNA Cleanup Kit (NEB; Art. No. T1030L), measured at a NanoDrop™ spectrophotometer (Thermo Fisher Scientific™; Art. No. ND-2000C). All reactions were conducted according to the manufacturer's manual. All resulting plasmids were verified by sequencing ("Ready2Run" by LGC Genomics GmbH; Berlin, Germany). The plasmids can be retrieved in GenBank format from the additional files (Additional file 4–Additional file 7). A summary of all plasmids that have been constructed during this study and their main elements is contained in Additional file 1: Table S2.

For construction of the *Agrobacterium tumefaciens* transformation plasmid pAt-DNG02774, plasmid pCBRS [91] and gDNA from *C. gorgonifer* strain NG\_p51 were taken as templates for PCR with Q5 polymerase. Q5 products Q5\_At\_02774\_1 to Q5\_At\_02774\_5 (Additional file 3: Table S1) were used for YRC. The correct assembly was verified by restriction digest of the resulting plasmid with MlI from Thermo Scientific. The plasmid contains the features needed for agrobacterium mediated transformation and an exchange cassette which replaces the first 2300 base-pairs of gene DNG\_02774 (i.e.: CgPKS4) with the hph marker.

For plasmid pCas9-DNG02774-sgRNA5 which was used for the Cas9 mediated mutation of DNG\_02774, the two precursor plasmids, pCas9-scaff and pCas9-hph were constructed. pCas9-scaff was assembled by using HiFi and PCR fragments Q5-Cas9-Scaff-1 to Q5-Cas9-Scaff-7 (Additional file 1: Table S1). Templates for the PCRs were the plasmid pFC331 [54], gDNA from an *Aspergillus fumigatus* wild-type isolate (in-house strain collection AIT-#1232), gDNA from *Aspergillus nidulans* WIM 126 [92], plasmid psgRNA [93], plasmid pCSN44 [94] and the plasmid pRS426 [95].

For the construction of pCas9-hph, plasmid pCas9-scaff was cut with Eco72I (Thermo Scientific™; Art. No. ER0361) and used for HiFi assembly with PCR fragment Q5-Cas9-hph that contains the hph marker (Additional file 1: Table S1).

For construction of plasmid pCas9-DNG02774-sgRNA5, plasmid pCas9-hph was cut with Bsp1407I (Thermo Scientific™; Art. No. ER0932), and assembled together with an sgRNA oligonucleotide with the sequence 5'-gtaccagacgaatctacacaCAAGGACTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTCTCCTTGACGAAGTAGCCACCgttttagagctagaatagc-3'. The sgRNA has the sequence: 5'-TCCTTGACGAAGTAGCCACC | GGG (PAM)-3' and anneals around base pair 275 downstream of the start codon of CgPKS4 (DNG\_02774).

### Strain construction

All fungal strains used during this study are listed in Additional file 1: Table S3.

### Agrobacterium transformation

Strain Cg-At-02774-1 was generated by disruption of CgPKS4 by agrobacterium mediated transformation with plasmid pAt-DNG02774. The transformation was conducted according to [96, 97] with the deviation that fungal cultures were pre-grown on CM-agar plates and that for co-cultivation of bacteria and fungi, not cellophane or agar-blocks were used, but 200 µL of a 1:1 mixture of a fungal spore solution ( $1 \times 10^6$  spores/mL in Tween-20 0.1%) and an induced *Agrobacterium* culture broth (*Agrobacterium tumefaciens* carrying pAt-DNG02774) were plated on 10 MoserIND petri dishes each. After 4 days of incubation at 20 °C, the plates were overlain with 15 mL of ~40 °C warm selection media (YES+hygromycinB+cefotaxim) and incubated for further 3 days at 30 °C before picking transformants. After purification via single spore isolation, the proper insertion of the replacement cassette was verified by Q5®-PCR with primer pair At-HDR\_Diag (Additional file 1: Table S1) which amplifies the region that spans from outside of the homology regions and across the whole insert. The proper integration was assessed via size comparison of amplicons of the recipient strain (i.e. NG\_p51) and the mutant on an agarose gel.

### Disruption of CgPKS4 by protoplast transformation (Cas9 and HDR)

For the disruption of gene DNG\_02774 by Cas9 (i.e. Cg-Cas9-02774-1) or by a linear fragment (HDR; i.e. Cg-HDR-02774-1), 10 µg plasmid pCas9-DNG02774-sgRNA5 respectively 72 µL of an un-purified Q5®-PCR product (i.e. "HDR-fragment" which is the exact same replacement cassette as within pAt-DNG02774; (Additional file 1: Table S1) was transformed via a modified protoplast transformation protocol into strain NG\_p51 as described below. In case of the Cas9 approach, successful disruption was verified by amplification of the

promoter and 5' end of the gene DNG\_02774 (~2 kb) with primer pair Cas9-Diag (Additional file 1: Table S1) which showed a 600 base-pair deletion, starting at the sgRNA annealing site. The deletion was further confirmed by sequencing with primer 5'-ATCACTCCCATAGCCGTCATCG3-3' at the LGC genomics sequencing facility (Berlin, Germany). In case of the HDR approach, the whole replacement cassette was amplified from outside of the homology regions, exactly as described above for the verification of the agrobacterium transformation (i.e. Cg-At-02774-1).

#### Modified protoplastation and transformation protocol for *C. gorgonifer*

For protoplast preparation, spores were harvested with Tween® 20 0.1% from a yeast extract sucrose (YES) agar plate and filtered through a glass wool tip. 6 mL of YES-broth in an epprouvette were inoculated with a spore density of  $4 \times 10^6$  spores per mL media and incubated for 18 h at 30 °C at 180 rpm. The culture was transferred into a falcon tube and centrifuged for 5 min at 4 °C and 2218 rcf in a 5810 R Eppendorf centrifuge with swing-out rotor. The supernatant was discarded and 1 mL of Lysing solution was prepared: 12.5 mg Driselase from Basidiomycetes (Sigma-Aldrich; Art. No. D9515), 1 mg Chitinase, from *Streptomyces griseus* (Sigma-Aldrich; Art. No. C6137), 25 mg Lysing enzymes from *Trichoderma harzianum* (Sigma-Aldrich; Art. No. L1412) in 1 mL of "OsmoStock" solution (40 mM/L Citric acid, 1.1 mM/L  $MgSO_4$ ; set pH to 5.9 with 3 mol/L NaOH). The Lysing solution was added to the mycelium. The mixture was resuspended and transferred into 2 mL tubes and put into an "IKA KS4000 I control" shaker at 30 °C at 110 rpm for 5 h. After visual check of protoplasts, the mixture was centrifuged at 4 °C for 10 min at 6000 rcf in an Eppendorf 5424R centrifuge and as much intermediate layer (clean protoplasts on top, protoplasts and cell debris at bottom) as possible was removed. Protoplasts count was determined in a counting chamber (Neubauer improved; PC72.1—Carl Roth, Karlsruhe).

For preparation of protoplast cryo-stocks, the solution was set to  $4 \times 10^8$  protoplasts per mL with OsmoStock solution and the same amount of 50% glycerol was added and mixed. Aliquots of 90 µL were then slowly frozen in 1.5 mL Eppendorf tubes within two styro-foam boxes at -80 °C. Each tube holds protoplasts for 3 transformations (i.e.:  $6 \times 10^6$  protoplasts per transformation).

For protoplast transformation, one protoplast cryo-stock was thawed on ice.  $6 \times 10^6$  protoplasts (i.e. 30 µL) were taken and adjusted to 200 µL with ice-cold TN1 solution (i.e. modified STC buffer solution: Sorbitol 0.8 mol/L,  $CaCl_2$  50 mM/L, Tris-HCl pH 8 10 mM/L) [98]. 10 µg plasmid DNA (i.e.:

pCas9-DNG02774-sgRNA5) or 72 µL of non-purified Q5-PCR product HDR-fragment (Additional file 1: Table S1) and 50 µL of ice-cold TN2 (=TN1 with PEG 4000 40% (w/v)) solution were added and mixed gently per pipetting. The whole mixture was left on ice. After 30 min, 1 mL of TN2 solution was added, gently mixed and left at room temperature for 5 min. The mixture was then added to 45 mL of ~40 °C Bottom Agarose (AMM [99] + 0.8 mol/L Sucrose + 10 mM/L  $NaNO_3$  + 0.8% w/v low melt agarose (Carl Roth; Art. No. 6351.2) and equally distributed into 3 petri dishes. After 2 h at room temperature, 15 mL of 45 °C Top agarose (same recipe as Bottom Agarose but 1.5% (w/v) low melt agarose and 0.05 mg/mL hygromycin B (Merck Millipore; Art. No. US1400052; potency: 1140 µg/mg) was added to each plate. After solidification, plates were incubated at 30 °C until colonies appeared on the surface (~7 days). Colonies were picked and two rounds of single spore isolation (strike out—pick—grow until sporulation) were conducted on Aspergillus complete media agar plates (CM by DR. C. F. Robert [99]) with hygromycinB before preparing spore cryo-stocks.

#### Spore-cryo-stock preparation

The fungus was grown on a CM plate and incubated for 6 days at 30 °C. A spore suspension with Tween-20 0.1% (v/v) was prepared, mixed with a 50% (w/v) glycerol solution at a ratio of 1:1 and  $5 \times 750$  µL aliquots were frozen at -80 °C.

#### Fermentation and extraction for rasfonin standard preparation

Fermentation conditions, isolation and structural elucidation was carried out according to Labuda et al. [100]. The fungal spore suspension ( $5.0 \times 10^6$  spores/mL) was obtained after 7 days cultivation of the fungus (*Cephalotrichum gorgonifer* NG\_p51) on a potato dextrose agar (PDA, VWR Chemicals, Austria). Five colony plugs were cut (each ca 1×1 cm) and thoroughly mixed (on vortex for 2 min) with 30 mL of physiological solution (0.9% NaCl) in a sterile, 50 mL Falcon tube. In total, 2 kg of par-boiled rice was used as a production substrate. Briefly, a total of 100 g of rice mixed with 100 mL deionized water was let soaked for ca one hour in 1000 mL capacity Erlenmeyer flasks, and then sterilized under standard autoclaving conditions (20 min, 121 °C, under pressure). Each flask (n=20) was inoculated with 100 µL of spore suspension at the central part of the surface. The flasks were cultivated in perforated plastic bags for 28 days at 30 °C in the dark. At the end of the cultivation, the rice cultures were checked for purity, cut into small pieces and the whole content of the plates (fungal colonies with rice-substrate) was harvested into a 5 L glass flask. No heating

deactivation nor drying was performed. The material was then mixed with 2 L of acetonitrile (ACN, Honeywell, Seelze, Germany). After vigorous stirring for 2 min in three subsequent steps (with ca 20 min in between), the mixture was filtered through a steel sieve in order to separate the solid particles (fungus and medium). The remaining residual water (generated by condensation of the water on plates during fungal growth) was removed by the addition of 10 g of anhydrous sodium sulphate. The organic phase was then filtered through a filter paper (270 mm i.d., Macherey–Nagel, Düren, Germany) and concentrated under reduced pressure at 45 °C (Büchi Rotavapor R-114, Flawil, Switzerland). The whole extraction procedure was repeated twice and yielded 3.5 g of crude culture extract.

#### Isolation of secondary metabolites for rasfonin standard preparation

The crude extract was purified by reversed-phase silica gel vacuum flash chromatography (Interchim, puriFlash®450, Montluçon Cedex, France), using three consecutive Interchim puriFlash® 32 g silica IR-50C18-F0025 flash columns (particle size: 50 µm). The columns were eluted with a binary solvent gradient (solvent A: H<sub>2</sub>O, solvent B: ACN). The starting linear gradient from 10% B to 27% B during 25 min at a flow rate 15 mL/min was followed by an isocratic gradient at 52% B for 10 min. Then a linear gradient from 52 to 66% B over 7 min was applied at the same flow rate and finally the column was washed starting with 100% B for 10 min followed by 100% A for 10 min at a flow rate 15–30 mL/min. UV 254 nm and UV scan 200–400 nm mode were used for detection and final separation of 7 main fractions (F1–F7), which were consequently concentrated under reduced pressure at 45 °C. The target compound was found in fraction F5 (23–27 R<sub>f</sub>, yield: 370 mg). It was resolved in a solvent mix (1:1:1; ACN/CH<sub>3</sub>OH/H<sub>2</sub>O) and further purified by an Agilent 1260 Infinity preparative HPLC (USA) on a reversed phase column Gemini NX C-18 (21.20×150 mm, 5 µm, 110 Å). Gradient starting with 30% ACN and 70% H<sub>2</sub>O up to 90% ACN in 10 min (total time 34 min) and a flow rate of 25 mL/min. Four fractions (pF1–pF4, time slice each 0.5 min) were collected, of which pF4 contained the target-rasfonin. Yield of rasfonin in pF4 (t<sub>R</sub> 17.85–18.70 min) after one stage of prep HPLC was 250 mg. For purity check an Agilent 1200 system was used with the same stationary phase (Gemini 5 µm NX-C18 110 Å, 150×2 mm) and gradient program at a flow rate 0.3 mL/min. Analytical chromatogram of purified rasfonin and its UV spectrum is presented on Additional file 1: Fig. S3. The identity of rasfonin was furthermore verified by NMR analysis (see below).

#### LC-HRMS

LC-HRMS measurements were carried out with a Vanquish UHPLC system coupled to a QExactive HF Orbitrap mass spectrometer (Thermo Fisher Scientific). For chromatographic separation a C18 reversed phase column (X-Bridge, 150×2.1 mm id, 3.5 µm particle size, Waters) was used with gradient elution applying H<sub>2</sub>O + 0.1% FA (eluent A) and methanol + 0.1% FA (eluent B) as eluents with a constant flow rate of 0.25 mL/min. After an initial minute with constant 10% eluent B, a linear increase towards 100% B was applied in 9 min followed by 3 constant minutes of 100% B. 7 min of re-equilibration with 10% B completed the method with a total length of 20 min. Ionisation was carried out in fast polarity switching mode and full scan mass spectra were recorded with a mass range of *m/z* 100–1000 and a resolution of 120 000 at *m/z* 200 (FWHM). MS/MS measurements for [M+H]<sup>+</sup> of Rasfonin were carried out in a data dependant method using an inclusion list. The selected ions were fragmented with stepped normalized collision energy (25, 35, 45 eV) and recorded with a resolution of 15 000 at *m/z* 200 (FWHM) at mass ranges *m/z* 50–460. For compound identification the in-house purified reference standard of Rasfonin (see above) was dissolved and diluted to a concentration of 1 mg/L and measured with the described method within the same sequence as all biological samples.

#### NMR

All NMR spectra were recorded on a Bruker Avance II 400 (resonance frequencies 400.13 MHz for <sup>1</sup>H and 100.63 MHz for <sup>13</sup>C) equipped with a 5 mm N<sub>2</sub>-cooled cryoprobe (Prodigy) with z-gradients at room temperature with standard Bruker pulse programs. The sample was dissolved in 0.6 mL of CDCl<sub>3</sub> (99.8% D). Chemical shifts are given in ppm, referenced to residual solvent signals (7.26 ppm for <sup>1</sup>H, 77.0 ppm for <sup>13</sup>C). <sup>1</sup>H NMR data were collected with 32 k complex data points and apodised with a Gaussian window function (lb = −0.3 Hz and gb = 0.3 Hz) prior to Fourier transformation. <sup>13</sup>C spectrum with WALTZ16 <sup>1</sup>H decoupling was acquired using 64 k data points. Signal-to-noise enhancement was achieved by multiplication of the FID with an exponential window function (lb = 1 Hz). All two-dimensional experiments were performed with 1 k×256 data points, while the number of transients (2–16 scans) and the sweep widths were optimized individually. HSQC experiment was acquired using adiabatic pulse for inversion of <sup>13</sup>C and GARP-sequence for broadband <sup>13</sup>C-decoupling, optimized for <sup>1</sup>J(CH) = 145 Hz. For the NOESY spectrum a mixing time of 0.5 s was used.

### Structure determination of rasfonin (Fig. 3) standard by NMR

The isolated compound showed a  $[M+H]^+$  peak at  $m/z=435.2697$  1229 (calcd 435.2741 for  $C_{25}H_{39}O_6$ ) in the ESI spectrum which is in accordance with the expected molecular formula of  $C_{25}H_{38}O_6$ . The  $^1H$  and  $^{13}C$  nmr spectra revealed the presence of three doublet methyl groups at  $\delta_H/\delta_C$  1.54/13.33, 1.15/18.85, 0.78/20.59 ppm and two singlet methyl groups at 1.52/15.53, and 1.83/12.59 ppm, respectively. In addition, 6 olefinic protons could be detected. Two pairs of them are part from delocalized double bond systems due to conjugation to carboxylic carbons C-1 and C-1': H-2/C-2 and H-3/C-3 showed signals at  $\delta_H/\delta_C$  6.21/124.93 and 7.04/140.58 ppm, whereas the system 2' and 3' resonated at  $\delta_H/\delta_C$  5.81/115.01 and 7.34/150.96 ppm, respectively. Furthermore, two oxygenated methines and two oxygenated methylene groups were identified in the nmr spectra. Detailed analysis of all two-dimensional nmr spectra proved the presence of rasfonin for the isolated compound. All assigned chemical shifts (Additional file 1: Table S4, Figs. S4–S9) are in agreement with those already reported for rasfonin by Fujimoto et al. [101].

### Rasfonin analysis of *CgPKS4* knock-out strains

For rasfonin analysis of *CgPKS4*-mutants, triplicates of each knock-out strain and the wildtype NG\_p51 were cultivated. Fungal spores were plated on YES-agar plates supplemented with 100 mM  $NO_3$  and cultivated at 30 °C for 7 days. The culture+agar was then cut into ~1 cm<sup>2</sup> pieces, frozen in liquid nitrogen and lyophilised in a Christ Alpha 2–4 LSCbasic lyophilizer. The freeze-dried material was ground in a Retsch MIXER MILL MM 400 ball mill in a liquid nitrogen cooled 50 mL grinding jar with one 20 mm stainless steel grinding ball per jar at 25 Hz for 60 s. 0.1 g of the ground material were extracted with 1.5 mL ACN with 0.1% FA acid on an IKA Vibrax VXR basic shaker at 1500 rpm for 1 h at room temperature. The slurry was then centrifuged for 20 min at 20,238 rcf and at room temperature in an Eppendorf 5424 centrifuge. The supernatant was then filtrated through a 0.22  $\mu$ M PTFE syringe filter, and the extract was reduced in a prewarmed (45 °C) Eppendorf concentrator plus for 30 min at V-HV settings. The reduced extracts were refilled to 300  $\mu$ L with ACN and filtered again to remove precipitate that formed during concentration. The extracts were directly measured in via HPLC and LC-HRMS.

### Measurement of knock-out-mutants via HPLC

Samples were analyzed with an Agilent 1200 system with a reversed phase column Phenomenex Gemini NX C-18

(2×150 mm, 5  $\mu$ m, 110 Å). Injection volume of samples was 20  $\mu$ L. The flow rate was 0.3 mL/min with a binary gradient of A from 15% (ACN+0.1% FA) to 95% in the time interval 2–31 min (Solvent B: MQ-H<sub>2</sub>O+0.1% FA). Rasfonin was detected with an Agilent G1315C DAD detector at 280 nm. Elution time of rasfonin was at around 21.1 min. For reference measurement and sample spiking, a 25  $\mu$ g/mL rasfonin standard produced by our laboratory (see above) was used.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40694-023-00158-x>.

**Additional file 1. Table S1:** Information about oligonucleotides, strain verification, LC-HRMS analysis of KO strains, plasmid information, strain information, UV spectrum of rasfonin, nmr data of rasfonin, Growth of *C. gorgonifer* at different temperatures.

**Additional file 2.** Transcriptome dataset and analysis of *C. gorgonifer* NG\_p51 at 24h and 48h in AMM liquid media.

**Additional file 3.** Genbank file with the rasfonin BGC and surrounding genetic locus, information on *CgPKS4* disruptions of KO strains.

**Additional file 4.** Plasmid map pAt\_DNG02774 in genbank format.

**Additional file 5.** Plasmid map pCas9-DNG02774-sgRNA5 in genbank format.

**Additional file 6.** Plasmid map pCas9-hph in genbank format.

**Additional file 7.** Plasmid map pCas9-scaff in genbank format.

### Acknowledgements

We want to thank colleagues from the Mortensen laboratory (Technical University of Denmark, Søtofts Plads, Kongens Lyngby, Denmark) that provided the plasmid templates (pFC331) for the Cas9 gene and the AMA1 region. We are also thankful to Dragana Bandian (AIT, Austrian Institute of Technology) who provided the gDNAs of fungal wild-type isolates for plasmid construction.

### Author contributions

AS (study design, protocol establishment, micro- and molecular biology, figures, writing), LS-R (Biosynthesis pathway of rasfonin, linking of rasfonin to BGC, study design, writing, figures), HB (Transcriptome analysis, BGC prediction), LS (chemical analysis, preliminary data acquisition), RL (Species description, morphology, rasfonin standard preparation and verification, figures), UG (Genome assembly, Gene annotation), MG (Agrobacterium Trafo, Isolation, strain information and identification), MB (NMR analysis), MD (LC-HRMS chemical analysis), EG (HPLC rasfonin method establishment and detection), AG (HPLC-chemical analysis), MS (screening, chemical analysis), JS (study design, manuscript writing and revision, funding acquisition). All authors read and approved the final manuscript.

### Funding

Open access funding provided by Austrian Science Fund (FWF). Open access publishing is funded by the Austrian Science Fund (FWF). Work was funded by Grant LS19-009 "TASSMATA" and "Bioactive Microbial Metabolites" Grant K3-G-2/08-2020 from the NFB Lower Austria Science Fund and Grant P32790 "ChroCosm" from the FWF Austrian Science Fund to JS. LRS is a Hertha-Firnberg fellow supported by grant T 1266-B from the FWF.

### Availability of data and materials

The datasets that were generated during this study are accessible at following repositories: Genome and Gene annotation of *C. gorgonifer* NG\_p51: European Nucleotide Archive (ENA) at EMBL-EBI. Project number PRJEB15373 which can be accessed at <http://www.ebi.ac.uk/ena/data/view/ONZQ01000001-ONZQ01000048> [40]. Transcriptome data of *C. gorgonifer* NG\_p51 at 24 h and 48 h: The data discussed in this publication have been deposited in NCBI's Gene

Expression Omnibus [46] and are accessible through GEO Series accession number GSE217303 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE217303>). The datasets supporting the conclusions of this article are included within the article and its additional files.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>Department of Applied Genetics and Cell Biology, Institute of Microbial Genetics, University of Natural Resources and Life Sciences, Vienna (BOKU), Campus Tulln, Konrad Lorenz Strasse 24, 3430 Tulln an der Donau, Austria. <sup>2</sup>Research Platform Bioactive Microbial Metabolites (BiMM), Konrad Lorenz Strasse 24, 3430 Tulln an der Donau, Austria. <sup>3</sup>Department for Farm Animals and Veterinary Public Health, Institute of Food Safety, Food Technology and Veterinary Public Health, Unit of Food Microbiology, University of Veterinary Medicine Vienna, Veterinärplatz 1, 1210 Vienna, Austria. <sup>4</sup>Department of Bioinformatics, Technical University of Munich, TUM School of Life Sciences Weihenstephan, Freising, Germany. <sup>5</sup>AIT Austrian Institute of Technology GmbH, Bioresources, 3430 Tulln, Austria. <sup>6</sup>Department of Chemistry, Institute of Chemistry of Renewable Resources, University of Natural Resources and Life Sciences Vienna (BOKU), Konrad-Lorenz-Straße 24, 3430 Tulln, Austria. <sup>7</sup>Department of Agrobiotechnology (IFA-Tulln), Institute of Bioanalytics and Agro-Metabolomics, University of Natural Resources and Life Sciences, Vienna (BOKU), Konrad Lorenz Strasse 20, 3430 Tulln an der Donau, Austria. <sup>8</sup>Core Facility Bioactive Molecules, Screening and Analysis, University of Natural Resources and Life Sciences, Vienna, 3430 Tulln an der Donau, Austria. <sup>9</sup>Present Address: DGforLife, Operations – Research and Development, Via Albert Einstein, Marcallo c.C., 20010 Milan, Italy. <sup>10</sup>Present Address: German Heart Center Munich, Technical University Munich, Lazarettstraße 36, 80636 Munich, Germany. <sup>11</sup>Present Address: Department of Biological Chemistry, Faculty of Chemistry, University of Vienna, Josef-Holaubek-Platz 2, 1090 Vienna, Austria.

Received: 8 November 2022 Accepted: 28 April 2023

Published online: 24 June 2023

## References

- Liu Y-N, Zhang T-J, Lu X-X, Ma B-L, Ren A, Shi L, et al. Membrane fluidity is involved in the regulation of heat stress induced secondary metabolism in *Ganoderma lucidum*. *Environ Microbiol*. 2017;19(4):1653–68.
- Overy D, Correa H, Roullier C, Chi WC, Pang KL, Rateb M, et al. Does osmotic stress affect natural product expression in fungi? *Mar Drugs*. 2017;15(8):1.
- Wong HJ, Mohamad-Fauzi N, Rizman-Idid M, Convey P, Alias SA. Protective mechanisms and responses of micro-fungi towards ultraviolet-induced cellular damage. *Polar Sci*. 2019;20:19–34.
- Scherlach K, Sarkar A, Schroeckh V, Dahse H-M, Roth M, Brakhage AA, et al. Two induced fungal polyketide pathways converge into antiproliferative spiroanthrones. *ChemBioChem*. 2011;12(12):1836–9.
- Zehetbauer F, Seidl A, Berger H, Sulyok M, Kastner F, Strauss J. RimO (SrrB) is required for carbon starvation signaling and production of secondary metabolites in *Aspergillus nidulans*. *Fungal Genet Biol*. 2022;162:103726.
- Wang G, Ran H, Fan J, Keller NP, Liu Z, Wu F, et al. Fungal-fungal cocultivation leads to widespread secondary metabolite alteration requiring the partial loss-of-function VeA1 protein. *Sci Adv*. 2022;8(17):eabo6094.
- Bazafkan H, Dattenböck C, Böhmendorfer S, Tisch D, Stappeler E, Schmoll M. Mating type-dependent partner sensing as mediated by VEL1 in *Trichoderma reesei*. *Mol Microbiol*. 2015;96(6):1103–18.
- Pusztahelyi T, Holb I, Pócsi I. Secondary metabolites in fungus-plant interactions. *Front Plant Sci*. 2015;6:1.
- Alarcón J, Águila S. Lovastatin production by *Pleurotus ostreatus*: effects of the C:N ratio. *Zeitschrift für Naturforschung C*. 2006;61(1–2):95–8.
- Ziemons S, Koutsantas K, Becker K, Dahlmann T, Kück U. Penicillin production in industrial strain *Penicillium chrysogenum* P2niaD18 is not dependent on the copy number of biosynthesis genes. *BMC Biotechnol*. 2017;17(1):16.
- Atongbiik Achaglinkame M, Opoku N, Amagloh FK. Aflatoxin contamination in cereals and legumes to reconsider usage as complementary food ingredients for Ghanaian infants: a review. *J Nutr Intermed Metab*. 2017;10:1–7.
- Farhadi A, Fakhri Y, Vasseghian Y, Huseyn E, Mousavi KA. Prevalence and concentration of fumonisins in cereal-based foods: a global systematic review and meta-analysis study. *Environ Sci Pollut Res Int*. 2021;28(17):20998–1008.
- Giorni P, Bertuzzi T, Battilani P. Impact of fungi co-occurrence on mycotoxin contamination in maize during the growing season. *Front Microbiol*. 2019;10:1.
- Keller NP. Fungal secondary metabolism: regulation, function and drug discovery. *Nat Rev Microbiol*. 2019;17(3):167–80.
- Sørensen T, Petersen C, Fechte LI, Nielsen KL, Sondergaard TE. A highly contiguous genome assembly of *arhtrium puccinoides*. *Genome Biol Evol*. 2022;14(1):evac010.
- Graham-Taylor C, Kamphuis LG, Derbyshire MC. A detailed in silico analysis of secondary metabolite biosynthesis clusters in the genome of the broad host range plant pathogenic fungus *Sclerotinia sclerotiorum*. *BMC Genomics*. 2020;21(1):7.
- Blackwell M. The fungi: 1, 2, 3 ... 5.1 million species? *Am J Bot*. 2011;98(3):426–38.
- Gorfer M, Blumhoff M, Klaubauf S, Urban A, Inselsbacher E, Bandian D, et al. Community profiling and gene expression of fungal assimilatory nitrate reductases in agricultural soils. *ISME J*. 2011;5(11):1771–83.
- Klaubauf S, Inselsbacher E, Zechmeister-Boltenstern S, Wanek W, Gottsberger R, Strauss J, et al. Molecular diversity of fungal communities in agricultural soils from Lower Austria. *Fungal Divers*. 2010;44(1):65–75.
- Gorfer M, Klaubauf S, Berger H, Strauss J. The fungal contribution to the nitrogen cycle in agricultural soils. 2014. p. 209–25.
- Reyes-Dominguez Y, Bok JW, Berger H, Shwab EK, Basheer A, Gallmetzer A, et al. Heterochromatic marks are associated with the repression of secondary metabolism clusters in *Aspergillus nidulans*. *Mol Microbiol*. 2010;76(6):1376–86.
- Collemare J, Seidl MF. Chromatin-dependent regulation of secondary metabolite biosynthesis in fungi: is the picture complete? *FEMS Microbiol Rev*. 2019;43(6):591–607.
- Gacek A, Strauss J. The chromatin code of fungal secondary metabolite gene clusters. *Appl Microbiol Biotechnol*. 2012;95(6):1389–404.
- Pfannenstiel B, Keller N. On top of biosynthetic gene clusters: How epigenetic machinery influences secondary metabolism in fungi. *Biotechnol Adv*. 2019;37:1.
- Schüller A, Studt-Reinhold L, Strauss J. How to completely squeeze a Fungus—advanced genome mining tools for novel bioactive substances. *Pharmaceutics*. 2022;14(9):1837.
- Zutz C, Bacher M, Parich A, Kluger B, Gacek-Matthews A, Schuhmacher R, et al. Valproic acid induces antimicrobial compound production in *doratomyces microspores*. *Front Microbiol*. 2016;7:1.
- Bachleitner S, Sørensen JL, Gacek-Matthews A, Sulyok M, Studt L, Strauss J. Evidence of a demethylase-independent role for the H3K4-specific histone demethylases in *Aspergillus nidulans* and *Fusarium graminearum* secondary metabolism. *Front Microbiol*. 2019;10:1759.
- Gacek-Matthews A, Berger H, Sasaki T, Wittstein K, Gruber C, Lewis ZA, et al. KdmB, a jumonji histone H3 demethylase, regulates genome-wide H3K4 trimethylation and is required for normal induction of secondary metabolism in *Aspergillus nidulans*. *PLoS Genet*. 2016;12(8):e1006222.
- Niehaus E-M, Rindermann L, Janevska S, Münsterkötter M, Güldener U, Tudzynski B. Analysis of the global regulator Lae1 uncovers a connection between Lae1 and the histone acetyltransferase HAT1 in *Fusarium fujikuroi*. *Appl Microbiol Biotechnol*. 2018;102(1):279–95.

30. Atanasoff-Kardjaleff AK, Studt L. Secondary metabolite gene regulation in mycotoxigenic fusarium species: a focus on chromatin. *Toxins* (Basel). 2022;14(2):1.
31. Sulyok M, Stadler D, Steiner D, Krska R. Validation of an LC-MS/MS-based dilute-and-shoot approach for the quantification of > 500 mycotoxins and other secondary metabolites in food crops: challenges and solutions. *Anal Bioanal Chem*. 2020;412(11):2607–20.
32. Hou B, Liu S, Li E, Jiang X. Different role of raptor and rictor in regulating rasfonin-induced autophagy and apoptosis through the glycolytic pathway in renal carcinoma cells. *Chem Biodivers*. 2020;17(12):e2000743.
33. Lu Q, Yan S, Sun H, Wang W, Li Y, Yang X, et al. Akt inhibition attenuates rasfonin-induced autophagy and apoptosis through the glycolytic pathway in renal cancer cells. *Cell Death Dis*. 2015;6(12):e2005.
34. Xiao Z, Li L, Li Y, Zhou W, Cheng J, Liu F, et al. Rasfonin, a novel 2-pyrone derivative, induces ras-mutated Panc-1 pancreatic tumor cell death in nude mice. *Cell Death Dis*. 2014;5(5):e1241-e.
35. Zhang F, Yan TQ, Guo W. Rasfonin inhibits proliferation and migration of osteosarcoma 143B cells. *Beijing Da Xue Xue Bao Yi Xue Ban*. 2019;51(2):234–8.
36. Sandoval-Denis M, Guarro J, Cano-Lira JF, Sutton DA, Wiederhold NP, de Hoog GS, et al. Phylogeny and taxonomic revision of Microasaceae with emphasis on synnematosus fungi. *Stud Mycol*. 2016;83(1):193–233.
37. Woudenberg JHC, Sandoval-Denis M, Houbraken J, Seifert KA, Samson RA. Cephalotrichum and related synnematosus fungi with notes on species from the built environment. *Stud Mycol*. 2017;88(1):137–59.
38. Domsch KH, Gams W, Anderson TH. Compendium of soil fungi. Volume 1. London: Academic Press (London) Ltd.; 1980. viii + 860pp. p.
39. Ellis MB. Dematiaceous Hyphomycetes: Kew, Commonwealth Mycological Institute; 1971. 608 pp. p.
40. Genome sequence data and gene annotation of Cephalotrichum gorgonifer NG\_p51: European Nucleotide Archive; 2022. Available from: <http://www.ebi.ac.uk/ena/data/view/ONZQ01000001-ONZQ01000048>.
41. Westphal KR, Muurmann AT, Paulsen IE, Nørgaard KTH, Overgaard ML, Dall SM, et al. Who needs neighbors? PKS8 is a stand-alone gene in *Fusarium graminearum* responsible for production of gibberpyrones and propilpyrone B. *Molecules*. 2018;23(9):1.
42. Brown DW, Yu JH, Kelkar HS, Fernandes M, Nesbitt TC, Keller NP, et al. Twenty-five coregulated transcripts define a sterigmatocystin gene cluster in *Aspergillus nidulans*. *Proc Natl Acad Sci USA*. 1996;93(4):1418–22.
43. Medema MH, Kottmann R, Yilmaz P, Cummings M, Biggins JB, Blin K, et al. Minimum information about a biosynthetic gene cluster. *Nat Chem Biol*. 2015;11(9):625–31.
44. Robey MT, Caesar LK, Drott MT, Keller NP, Kelleher NL. An interpreted atlas of biosynthetic gene clusters from 1000 fungal genomes. *Proc Natl Acad Sci*. 2021;118(19):e202030118.
45. Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema Marnix H, et al. antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucl Acids Res*. 2021;49(W1):W29–W35.
46. Edgar R, Domrachev M, Lash AE. Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucl Acids Res*. 2002;30(1):207–10.
47. Zhou ZZ, Zhu HJ, Lin LP, Zhang X, Ge HM, Jiao RH, et al. Dalmanol biosyntheses require coupling of two separate polyketide gene clusters. *Chem Sci*. 2019;10(1):73–82.
48. Ugai T, Minami A, Fujii R, Tanaka M, Oguri H, Gomi K, et al. Heterologous expression of highly reducing polyketide synthase involved in betaenone biosynthesis. *Chem Commun (Camb)*. 2015;51(10):1878–81.
49. Bonsch B, Belt V, Bartel C, Duensing N, Koziol M, Lazarus CM, et al. Identification of genes encoding squalenstatin S1 biosynthesis and in vitro production of new squalenstatin analogues. *Chem Commun*. 2016;52(41):6777–80.
50. Trenti F, Cox RJ. Structural revision and biosynthesis of the fungal phyto-toxins phyllostictines A and B. *J Nat Prod*. 2017;80(5):1235–40.
51. Chen LH, Lin CH, Chung KR. A nonribosomal peptide synthetase mediates siderophore production and virulence in the citrus fungal pathogen *Alternaria alternata*. *Mol Plant Pathol*. 2013;14(5):497–505.
52. Qiao K, Chooi Y-H, Tang Y. Identification and engineering of the cytochalasin gene cluster from *Aspergillus clavatus* NRRL 1. *Metab Eng*. 2011;13(6):723–32.
53. Li D, Tang Y, Lin J, Cai W. Methods for genetic transformation of filamentous fungi. *Microb Cell Fact*. 2017;16(1):168.
54. Nødvig CS, Nielsen JB, Kogle ME, Mortensen UH. A CRISPR-Cas9 system for genetic engineering of filamentous fungi. *PLoS ONE*. 2015;10(7):e0133085.
55. Wu J-D, Chou J-C. Optimization of protoplast preparation and regeneration of a medicinal fungus *Antrodia cinnamomea*. *Mycobiology*. 2019;47(4):483–93.
56. Gaffoor I, Trail F. Characterization of two polyketide synthase genes involved in zearalenone biosynthesis in *Gibberella zeae*. *Appl Environ Microbiol*. 2006;72(3):1793–9.
57. Lonjon F, Rengel D, Roux F, Henry C, Turner M, Le Ru A, et al. HpaP Sequesters HrpJ, an essential component of *Ralstonia solanacearum* virulence that triggers necrosis in *Arabidopsis*. *Mol Plant Microbe Interact*. 2019;33(2):200–11.
58. Chiang Y-M, Szwedczyk E, Davidson AD, Keller N, Oakley BR, Wang CCC. A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. *J Am Chem Soc*. 2009;131(8):2965–70.
59. Dalmasis B, Schumacher J, Moraga J, Le Pêcheur P, Tudzynski B, Collado IG, et al. The *Botrytis cinerea* phytotoxin botcinic acid requires two polyketide synthases for production and has a redundant role in virulence with botrydial. *Mol Plant Pathol*. 2011;12(6):564–79.
60. Sutherland A, Auclair K, Vederas JC. Recent advances in the biosynthetic studies of lovastatin. *Curr Opin Drug Discov Devel*. 2001;4(2):229–36.
61. Abe Y, Suzuki T, Ono C, Iwamoto K, Hosobuchi M, Yoshikawa H. Molecular cloning and characterization of an ML-236B (compactin) biosynthetic gene cluster in *Penicillium citrinum*. *Mol Genet Genomics*. 2002;267(5):636–46.
62. Wang WG, Wang H, Du LQ, Li M, Chen L, Yu J, et al. Molecular basis for the biosynthesis of an unusual chain-fused polyketide, Gregatin A. *J Am Chem Soc*. 2020;142(18):8464–72.
63. Atanasoff-Kardjaleff AK, Seidl B, Steinert K, Daniliuc CG, Schuhmacher R, Humpf HU, et al. Biosynthesis of the isocoumarin derivatives fusamarens is mediated by the PKS8 gene cluster in *Fusarium*. *Chembiochem*. 2022:e202200342.
64. von Barga KW, Niehaus EM, Krug I, Bergander K, Würthwein EU, Tudzynski B, et al. Isolation and structure elucidation of fujikurins A–D: products of the PKS19 gene cluster in *Fusarium fujikuroi*. *J Nat Prod*. 2015;78(8):1809–15.
65. Wiemann P, Sieber CMK, von Barga KW, Studt L, Niehaus E-M, Espino JJ, et al. Deciphering the cryptic genome: genome-wide analyses of the rice pathogen *Fusarium fujikuroi* reveal complex regulation of secondary metabolism and novel metabolites. *PLoS Pathog*. 2013;9(6):e1003475.
66. Xu W, Chooi Y-H, Choi JW, Li S, Vederas JC, Da Silva NA, et al. LovG: the thioesterase required for dihydromonacolin L release and lovastatin nonaketide synthase turnover in lovastatin biosynthesis. *Angew Chem Int Ed*. 2013;52(25):6472–5.
67. Droce A, Saei W, Jørgensen SH, Wimmer R, Giese H, Wollenberg RD, et al. Functional analysis of the fusarielin biosynthetic gene cluster. *Molecules* [Internet]. 2016; 21(12):1.
68. Newman AG, Vagstad AL, Belecki K, Scheerer JR, Townsend CA. Analysis of the cercosporin polyketide synthase CTB1 reveals a new fungal thioesterase function. *Chem Commun*. 2012;48(96):11772–4.
69. Campbell CD, Vederas JC. Biosynthesis of lovastatin and related metabolites formed by fungal iterative PKS enzymes. *Biopolymers*. 2010;93(9):755–63.
70. Jiang Y, Wang D, Yao Y, Eubel H, Künzler P, Möller IM, et al. MULocDeep: a deep-learning framework for protein subcellular and suborganellar localization prediction with residue-level interpretation. *Comput Struct Biotechnol J*. 2021;19:4825–39.
71. Boenisch MJ, Broz KL, Purvine SO, Chrisler WB, Nicora CD, Connolly LR, et al. Structural reorganization of the fungal endoplasmic reticulum upon induction of mycotoxin biosynthesis. *Sci Rep*. 2017;7:44296.
72. Menke J, Weber J, Broz K, Kistler HC. Cellular development associated with induced mycotoxin synthesis in the filamentous fungus *Fusarium graminearum*. *PLoS ONE*. 2013;8(5):e63077.

73. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 2018;35(6):1547–9.
74. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl Acids Res*. 1994;22(22):4673–80.
75. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol*. 1993;10(3):512–26.
76. Samson RAHJTUFJ CAB. Food and indoor fungi 2019.
77. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics*. 2010;27(4):578–9.
78. 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.
79. Walter MC, Rattei T, Arnold R, Güldener U, Münsterkötter M, Nenova K, et al. PEDANT covers all complete RefSeq genomes. *Nucl Acids Res*. 2009;37(suppl\_1):D408–D11.
80. Vienna Biocenter Core Facilities. Available from: <https://www.viennabiocenter.org/vbcf/>.
81. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–20.
82. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2009;25(14):1754–60.
83. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. *Gigascience*. 2021;10(2):1.
84. Picard MarkDuplicates. Available from: <https://broadinstitute.github.io/picard/>.
85. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166–9.
86. Law CW, Chen Y, Shi W, Smyth GK. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014;15(2):R29.
87. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26(6):841–2.
88. Bedtools Genomcov. Available from: <https://bedtools.readthedocs.io/en/latest/content/tools/genomcov.html>.
89. Cenis JL. Rapid extraction of fungal DNA for PCR amplification. *Nucl Acids Res*. 1992;20(9):2380.
90. Schumacher J. Tools for Botrytis cinerea: new expression vectors make the gray mold fungus more accessible to cell biology approaches. *Fungal Genet Biol*. 2012;49(6):483–97.
91. Németh MZ, Li G, Seress D, Pintye A, Molnár O, Kovács GM, et al. What is the role of the nitrate reductase (euknr) gene in fungi that live in nitrate-free environments? A targeted gene knock-out study in *Ampelomyces mycoparasites*. *Fungal Biol*. 2021;125(11):905–13.
92. Mooney JL, Yager LN. Light is required for conidiation in *Aspergillus nidulans*. *Genes Dev*. 1990;4(9):1473–82.
93. Schüller A, Wolansky L, Berger H, Studt L, Gacek-Matthews A, Sulyok M, et al. A novel fungal gene regulation system based on inducible VPR-dCas9 and nucleosome map-guided sgRNA positioning. *Appl Microbiol Biotechnol*. 2020;104(22):9801–22.
94. Staben C, Jensen B, Singer M, Pollock J, Schechtman M, Kinsey J, et al. Use of a bacterial hygromycin B resistance gene as a dominant selectable marker in *Neurospora crassa* transformation. *Fungal Genet Rep*. 1989;36(1):1.
95. Christianson TW, Sikorski RS, Dante M, Shero JH, Hieter P. Multifunctional yeast high-copy-number shuttle vectors. *Gene*. 1992;110(1):119–22.
96. Hanif M, Pardo AG, Gorfer M, Raudaskoski M. T-DNA transfer and integration in the ectomycorrhizal fungus *Suillus bovinus* using hygromycin B as a selectable marker. *Curr Genet*. 2002;41(3):183–8.
97. Gorfer M, Klaubauf S, Bandian D, Strauss J. *Cadophora finlandia* and *Phialocephala fortinii*: agrobacterium-mediated transformation and functional GFP expression. *Mycol Res*. 2007;111(7):850–5.
98. He Z-M, Price MS, O'Brien GR, Georgianna DR, Payne GA. Improved protocols for functional analysis in the pathogenic fungus *Aspergillus flavus*. *BMC Microbiol*. 2007;7(1):104.
99. Barratt RW, Johnson GB, Ogata WN. Wild-type and mutant stocks of *Aspergillus nidulans*. *Genetics*. 1965;52(1):233–46.
100. Labuda R, Bacher M, Gratzl H, Doppler M, Parich A, Aufy M, et al. Luteapyrone, a novel γ-Pyrone Isolated from the filamentous fungus *Metapochonia lutea*. *Molecules*. 2021;26(21):6589.
101. Fujimoto H, Okamoto Y, Sone E, Maeda S, Akiyama K, Ishibashi M. Eleven new 2-pyrones from a Fungi imperfecti, *Trichurus terrophilus*, found in a screening study guided by immunomodulatory activity. *Chem Pharm Bull*. 2005;53(8):923–9.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

