

Cite this: *Chem. Sci.*, 2024, 15, 3349

All publication charges for this article have been paid for by the Royal Society of Chemistry

# Discovery and heterologous biosynthesis of glycosylated polyketide luteodienoside A reveals unprecedented glucinol-mediated product offloading by a fungal carnitine O-acyltransferase domain†

Amr A. Arishi,<sup>ab</sup> Zhuo Shang,<sup>ac</sup> Ernest Lacey,<sup>de</sup> Andrew Crombie,<sup>e</sup> Daniel Vuong,<sup>e</sup> Hang Li,<sup>af</sup> Joe Bracegirdle,<sup>a</sup> Peter Turner,<sup>g</sup> William Lewis,<sup>g</sup> Gavin R. Flematti,<sup>a</sup> Andrew M. Piggott<sup>bd</sup> and Yit-Heng Chooi<sup>da</sup>

Luteodienoside A is a novel glycosylated polyketide produced by the Australian fungus *Aspergillus luteorubrus* MST-FP2246, consisting of an unusual 1-O-β-D-glucopyranosyl-*myo*-inositol (glucinol) ester of 3-hydroxy-2,2,4-trimethylocta-4,6-dienoic acid. Mining the genome of *A. luteorubrus* identified a putative gene cluster for luteodienoside A biosynthesis (*ltb*), harbouring a highly reducing polyketide synthase (HR-PKS, LtbA) fused at its C-terminus to a carnitine O-acyltransferase (cAT) domain. Heterologous pathway reconstitution in *Aspergillus nidulans*, substrate feeding assays and gene truncation confirmed the identity of the *ltb* cluster and demonstrated that the cAT domain is essential for offloading luteodienoside A from the upstream HR-PKS. Unlike previously characterised cAT domains, the LtbA cAT domain uses glucinol as an offloading substrate to release the product from the HR-PKS. Furthermore, the PKS methyltransferase (MT) domain is capable of catalysing *gem*-dimethylation of the 3-hydroxy-2,2,4-trimethylocta-4,6-dienoic acid intermediate, without requiring reversible product release and recapture by the cAT domain. This study expands the repertoire of polyketide modifications known to be catalysed by cAT domains and highlights the potential of mining fungal genomes for this subclass of fungal PKSs to discover new structurally diverse secondary metabolites.

Received 23rd September 2023

Accepted 4th January 2024

DOI: 10.1039/d3sc05008d

rsc.li/chemical-science

## Introduction

Fungal glycosides are a large and diverse class of secondary metabolites with significant biological activities and promising pharmaceutical applications.<sup>1</sup> Among the fungal glycosides, glycosylated polyketides are particularly noteworthy for their

broad spectrum of bioactivities. Some examples include the antifungal burnettramic acids from *Aspergillus burnettii*,<sup>2</sup> the cytotoxic cladionol A from *Gliocladium* sp.,<sup>3</sup> and the antibacterial epicoccamide from *Epicoccum purpurascens*.<sup>4,5</sup> However, despite these diverse bioactivities, the biosynthesis of glycosylated polyketides is still not well understood in fungi.

Technological advances in next-generation DNA sequencing and bioinformatics tools, coupled with an expanding knowledge of fungal secondary metabolite biosynthetic pathways, have seen genome mining emerge as an effective approach for the discovery of novel bioactive molecules.<sup>6</sup> It is well known that biosynthetic genes are clustered together in fungal genomes. Reconstitution of these biosynthetic gene clusters (BGCs) in heterologous expression hosts allows the function of each individual gene in a biosynthetic pathway to be interrogated, making it a powerful approach to link genes to biosynthetic steps and chemical structures.<sup>7</sup> These genes can then serve as beacons to guide genome mining for structurally related molecules. Hence, we became interested in dissecting the genetic and biochemical basis for glycosylated polyketides in fungi, with a view to facilitating targeted genome mining for new members of this important group of metabolites.

<sup>a</sup>School of Molecular Sciences, The University of Western Australia, Perth, WA 6009, Australia. E-mail: yitheng.chooi@uwa.edu.au

<sup>b</sup>Department of Botany and Microbiology, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

<sup>c</sup>School of Pharmaceutical Sciences, Shandong University, Jinan, Shandong 250012, China

<sup>d</sup>School of Natural Sciences, Macquarie University, Sydney, NSW 2109, Australia. E-mail: andrew.piggott@mq.edu.au

<sup>e</sup>Microbial Screening Technologies Pty. Ltd, Smithfield, NSW 2164, Australia

<sup>f</sup>School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou 510006, China

<sup>g</sup>School of Chemistry, The University of Sydney, NSW 2006, Australia

† Electronic supplementary information (ESI) available: Details regarding cultivation, isolation, structure elucidation, bioassay, X-ray crystallography, molecular biology, phylogenetic analysis, NMR data and NMR spectra. CCDC 2295166. For ESI and crystallographic data in CIF or other electronic format see DOI: <https://doi.org/10.1039/d3sc05008d>

As part of our ongoing chemotaxonomic exploration of endemic Australian *Aspergilli*,<sup>2,8–10</sup> we isolated a novel species, *Aspergillus luteorubrus* MST-FP2246, from a soil sample collected at White Mountains National Park in northern Queensland, Australia.<sup>11</sup> Herein, we report the discovery and structural characterisation of luteodienoside A (**1**), an abundant polar metabolite produced by *A. luteorubrus* featuring a novel glucinol ester of a *gem*-dimethylated polyketide. We employed heterologous pathway reconstitution, polyketide synthase (PKS) domain truncation, substrate feeding experiments and characterisation of pathway intermediates to elucidate the biosynthetic pathway to **1**.

## Results and discussion

### Isolation and characterisation of luteodienoside A from *Aspergillus luteorubrus*

*A. luteorubrus* MST-FP2246 was cultivated on basmati rice (2.0 kg) for 21 days at 24 °C. The grains were then extracted with acetone and the extract partitioned between ethyl acetate and water. The aqueous partition was fractionated by both reversed phase and normal phase silica gel chromatography to give an enriched extract (1.9 g, ~60% pure). A sub-sample of the extract (40 mg) was subjected to size exclusion chromatography on Sephadex LH-20, then reversed phase HPLC on C<sub>18</sub> silica, yielding pure luteodienoside A (**1**) as a light-yellow solid (22 mg).

HRESI(+)-MS analysis of **1** revealed a dehydrated protonated molecule at  $m/z$  505.2274 [ $M - H_2O + H$ ]<sup>+</sup>, indicative of a molecular formula C<sub>23</sub>H<sub>38</sub>O<sub>13</sub> requiring five double bond equivalents (DBEs). The UV-vis spectrum of **1** (Fig. S6†) showed a single  $\lambda_{max}$  at 240 nm, characteristic of an isolated diene. Analysis of the NMR data for **1** in DMSO-*d*<sub>6</sub> (Table S5†) revealed resonances for four methyl groups ( $\delta_{C-8}$  18.2,  $\delta_{C-9}$  23.0,  $\delta_{C-10}$  19.4,  $\delta_{C-11}$  14.3), four olefinic carbons ( $\delta_{C-4}$  134.6,  $\delta_{C-5}$  127.5,  $\delta_{C-6}$  127.4,  $\delta_{C-7}$  128.8) and one carbonyl group ( $\delta_{C-1}$  175.7), along with twelve oxygenated methines. COSY correlations spanning from H-5 to H-8 and HMBC correlations from H-11 to C-3 and C-5, and from H<sub>3</sub>-9/10 to C-1, C-2 and C-3, permitted assignment of a 3-hydroxy-2,2,4-trimethylpent-4-enoyl moiety (Fig. 1b). A large  $^3J_{H-6,7}$  coupling constant (14.5 Hz) and a ROESY correlation between H-6 and H-11 indicated a 4*E*,6*E* geometry for the diene. Based on the chemical shifts of the remaining 12 carbons ( $\delta_C$  60–100) and two remaining DBEs, the presence of two hexose moieties in the structure of **1** was hypothesised. COSY correlations spanning H-1' ( $\delta_H$  4.41) to H<sub>2</sub>-6' ( $\delta_H$  3.28/3.47) defined the first hexose moiety, with an HMBC correlation from H-4' ( $\delta_H$  4.52) to C-1 ( $\delta_C$  175.7) confirming an ester linkage between C-1 and C-4'. COSY correlations spanning H-1'' to H-6'' defined the second hexose moiety, with the absence of an anomeric centre and methylene group indicative of a carbocyclic sugar (inositol). An HMBC correlation from H-1' ( $\delta_H$  4.41) to C-4'' ( $\delta_C$  85.3) confirmed a 1,4-glycosidic linkage between the two sugars, thus completing the planar structure of **1**. Finally, single crystal X-ray diffraction analysis (Fig. S1†) confirmed the absolute stereostructure of **1** as being the 1-*O*- $\beta$ -D-glucopyranosyl-*myo*-inositol (glucinol) ester of 3*S*-hydroxy-2,2,4-trimethylocta-4,6-dienoic acid (Fig. 1a).

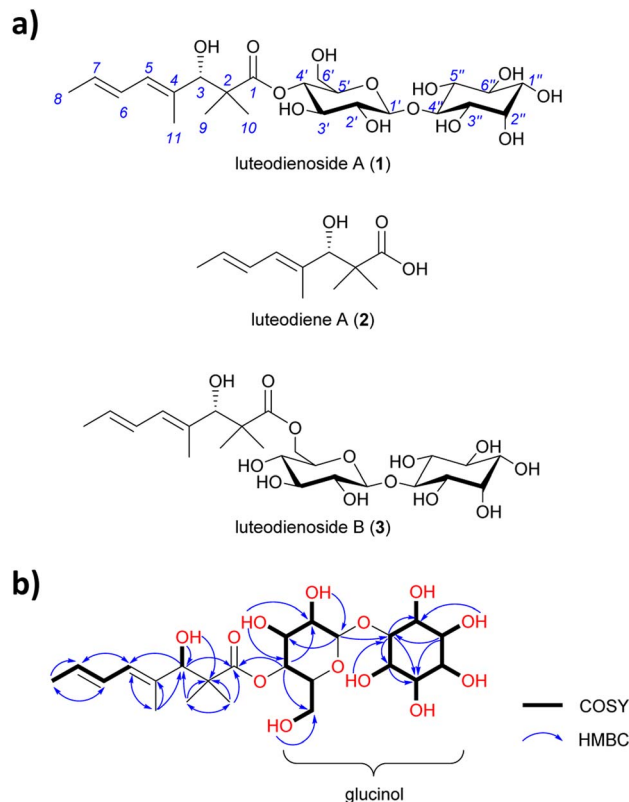


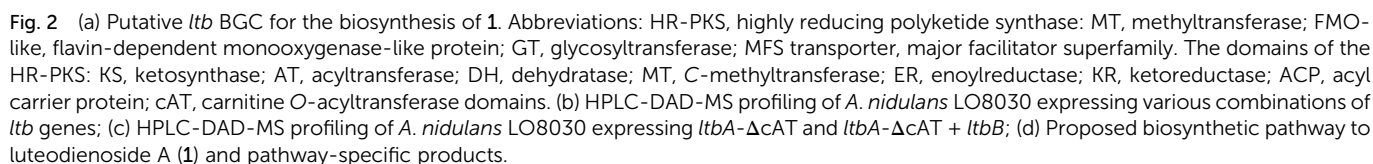
Fig. 1 (a) Structures of luteodienoside A (**1**), luteodiene A (**2**) and luteodienoside B (**3**) reported in this study. (b) Diagnostic 2D NMR correlations for **1**.

### Identification of luteodienoside A BGC featuring an HR-PKS harbouring a carnitine *O*-acyltransferase (CAT) domain

The unprecedented conjugation between glucinol and a polyketide chain in **1** prompted us to investigate the molecular basis for its biosynthesis. Retrobiosynthetic analysis of **1** suggested that the trimethylated polyketide chain is derived from acetyl-CoA, malonyl-CoA and *S*-adenosylmethionine (SAM), assembled by a highly reducing polyketide synthase (HR-PKS) containing a *C*-methyltransferase domain (MT). Additionally, the presence of glucinol suggested a requirement for one or more glycosyltransferase-encoding genes within the luteodienoside A BGC.

To identify the BGC for the biosynthesis of **1**, we sequenced the genome of *A. luteorubrus* using Illumina HiSeq 2500. Subsequently, genome assembly and gene annotation were performed using the automatic assembly for the fungi (AAFTF) pipeline and the funannotate pipeline, respectively.<sup>12,13</sup> The antibiotics and secondary metabolite analysis shell (antiSMASH) was used to predict BGCs within the genome of *A. luteorubrus*.<sup>14</sup> The analysis identified two BGCs harbouring co-localised HR-PKS and glycosyltransferase genes in scaffolds 4.1 and 9.1. Notably, the BGC located on scaffold 4.1 encodes a HR-PKS (denoted LtbA) with a domain architecture containing ketosynthase (KS), acyltransferase (AT), dehydratase (DH), *C*-methyltransferase (MT), enoylreductase (ER), ketoreductase (KR), acyl-carrier protein (ACP) and carnitine *O*-acyltransferase





(cAT) domains (Fig. 2a). HR-PKSs with C-terminal cAT domain fusions have been previously shown to produce polyketide chains with  $\alpha,\alpha$ -gem-dimethyl moieties,<sup>15</sup> as observed in the polyketide chain of **1**. This prompted us to focus on this BGC as the putative luteodienoside A BGC (*ltb*). In addition to *LtbA*, the enzymes encoded by the other genes within the *ltb* BGC include a glycosyltransferase-like family-2 protein (*LtbB*), a methyltransferase (*LtbC*), a flavin-dependent monooxygenase-like protein (*LtbD*) and a major facilitator superfamily (MFS) transporter (*LtbE*). The NCBI GenBank accession number for the *ltb* cluster is OR597289.

### Deciphering the biosynthetic pathway to luteodienoside A using heterologous expression reveals an unprecedented product offloading mechanism

After locating the candidate *ltb* BGC, we next sought to confirm its involvement in the biosynthesis of **1**. To this end, we reconstructed the whole BGC in the heterologous host *Aspergillus nidulans* LO8030, which has been genetically modified to minimise endogenous metabolite background.<sup>16</sup> The heterologous expression of the *ltb* BGC in *A. nidulans* was facilitated by the hybrid yeast-fungal artificial chromosome (pYFAC) expression system, which allows episomal expression of secondary metabolites in *A. nidulans*.<sup>17,18</sup> We first cloned the *ltbA* gene into the vector pYFAC-*pyrG* using yeast transformation-associated recombination in *Saccharomyces cerevisiae*. Subsequently, *ltbB*, *ltbC* and *ltbD*, both together and individually, were cloned into the vector pYFAC-*ribO*. Next, *A. nidulans* LO8030 protoplasts were introduced with pYFAC plasmids carrying the gene combination *ltbABCD* to express the entire *ltb* BGC. LC-MS analysis of the *n*-butanol extract from the transformant confirmed the production of **1**, along with two closely related compounds (**2** and **3**) that exhibited the same diene UV-vis profile as **1** (Fig. 2b and S6†). HRMS (Fig. S8†) and NMR (Table S6†) analysis of **2** identified the compound as the polyketide aglycone of **1**, which we named luteodiene A. Similarly, HRMS (Fig. S8†) and NMR (Table S5†) analysis of **3** identified the compound as a regioisomer of **1**, with the ester linkage between C-1 and C-6' rather than between C-1 and C-4'. Collectively, these results demonstrated that the proposed *ltb* BGC is responsible for the biosynthesis of **1**.

To characterise the biosynthetic logic of **1**, various gene combinations were heterologously expressed in *A. nidulans* (Fig. 2b). Expression of *ltbA* alone resulted in the detection of **2**, confirming it is the product of *LtbA*. Further EIC (+) *m/z* 505 analysis revealed that trace amount of **1** and **3** were present, but below the limit detectable by DAD (Fig. S9†). When *ltbABC* was expressed, peaks corresponding to **1**, **2** and **3** were observed with no significant difference in metabolite production levels compared to *ltbABCD*, suggesting that *ltbD*, encoding a truncated flavin-dependent monooxygenase-like protein (Fig. S23 and S24†), is not essential for the biosynthesis of **1**. To determine the roles of *ltbB* and *ltbC*, each gene was cloned both alone and co-expressed with *ltbA*. Expression of a transformant containing *ltbAC* resulted in the detection of carboxylic acid **2** by LC-MS analysis of a small-scale culture extract. Conversely,

when the transformant harbouring *ltbAB* was heterologously expressed, **1**, **2** and **3** were observed in the *n*-butanol extract. These results indicate that *ltbC*, despite encoding a methyltransferase containing the key active site residues (Fig. S22†), is not essential in the biosynthesis of **1**, and that *LtbA* and *LtbB* are minimally required for the biosynthesis of **1** (Fig. 2b).

Interestingly, **2** and **3** were not observed in cultures of the native producer *A. luteorubrus*. To investigate possible influence of the heterologous host *A. nidulans* on the biosynthesis of **1**, **2** and **3**, compound **1** was fed to the *A. nidulans* LO8030 host at a concentration of 10 mg L<sup>-1</sup>. Interestingly, after 12 h of incubation, **2** and **3** were both observed in the media, suggesting that *A. nidulans* produces endogenous enzymes capable of hydrolysing **1** to **2** and isomerising **1** to **3** (Fig. S3†). To corroborate these results, we conducted co-expression of intron-free *ltbA* and *ltbB* in *S. cerevisiae* BJ5464-NpgA. However, this attempt did not result in the detection of **1** or any pathway-specific products, possibly due to *LtbA* and/or *LtbB* not being functionally expressed in *S. cerevisiae*.

### Generation of a cAT-truncated version of *ltbA* identifies glucinol as an offloading substrate for the *LtbA* cAT domain

The biosynthesis of  $\alpha,\alpha$ -gem-dimethylated polyketides has previously been shown to involve HR-PKSs with C-terminal-fused cAT domains. In *Trichoderma virens*, the Tv6-931 cAT domain can release the immature monomethylated products then recapture them back on the ACP to the MT domain to perform a second methylation cycle to generate  $\alpha,\alpha$ -gem-dimethyl-containing products.<sup>15</sup> In contrast, the cAT domain of a HR-PKS from *Metarhizium anisopliae* (MaPKS17a) can catalyse the simple hydrolytic release of polyunsaturated polyketide products from the PKS.<sup>19</sup> Other cAT domains have been previously identified in the BGCs of AF-toxin and sordarin, but their functions remain unclear.<sup>20,21</sup> Although the final product of the *T. virens* BGC remains unknown, the cAT domain of the HR-PKS (Tv6-931) has been shown to utilise polyol nucleophiles to release the polyketide product as a polyol ester *in vitro*.<sup>15</sup>

To gain insights into the evolution and function of the *LtbA*-cAT domain, a phylogenetic analysis was constructed using amino acid sequences of 79 cAT domains extracted from various fungal HR-PKSs. The analysis revealed that the *LtbA*-cAT domain is distantly related to the known cAT domains (Fig. S2†), suggesting that it may have evolved a distinct function. We hypothesised that the *LtbA*-cAT domain may utilise glucinol as an offloading substrate to release the polyketide chain as a glucinol ester **1**. To test this hypothesis, we generated a truncated version of *LtbA* (KS-AT-DH-MT-ER-KR-ACP) lacking the region encoding the cAT domain. This modified *ltbA* construct was heterologously expressed both alone and in combination with *ltbB* in *A. nidulans*. When expressing *ltbA*- $\Delta$ cAT alone, a peak corresponding to **2** was detected (Fig. 2c), suggesting that PKS product **2** could be released from the *LtbA*-ACP by endogenous hydrolases of *A. nidulans* or hydrolysed non-enzymatically. Likewise, when co-expressing *ltbA*- $\Delta$ cAT with *ltbB*, the transformant only accumulates **2** and not **1**. These results suggest that (i) the *LtbA*-cAT domain is distinct from the





previously characterised Tv6-931 cAT domain in that it is not involved in recapturing the elongated polyketide acyl chain for dimethylation and (ii) the cAT domain is involved in offloading the  $\alpha,\alpha$ -gem-dimethylated polyketide acyl chain from the ACP domain, likely using glucinol as a nucleophile, to yield the glucinol ester **1** (Fig. 2c). Examination of the active site residues in LtbA cAT, Tv6-931 cAT, and MaPKS17a cAT domains, alongside the authentic cAT, revealed conservation of the catalytic residue histidine and the serine residue involved in coordinating water molecule within the active site (Fig. S17†).

The presence of gem-dimethyl groups is relatively rare in fungal polyketides, with only one assigned MT domain from Tv6-931 having been identified to date. The  $\alpha,\alpha$ -gem-dimethylation catalysed by the MT domain of Tv6-931 requires the cAT domain to recapture the immature  $\alpha$ -monomethyl polyketide acyl polyol ester chain for a second methylation cycle at the  $\alpha$  position then release it again as the  $\alpha,\alpha$ -gem-dimethyl polyketide polyol ester. In contrast, the *M. anisopliae* MaPKS17a only produces monomethylated products, indicating that the MT domain lacks the ability to perform dimethylation. In our study, the product of *ltbA*- $\Delta$ cAT (*i.e.* **2**) has already undergone  $\alpha,\alpha$ -gem-dimethylation by the MT domain (Fig. S8†). This is significant as the absence of the cAT domain did not affect gem-dimethylation, indicating that the LtbA-MT domain can doubly methylate the  $\alpha$  positions of the fully elongated polyketide acyl chain tethered on the ACP prior to its release, without being recaptured by the cAT domain (Fig. S16†).

To the best of our knowledge, this is the first report of a fungal cAT domain utilising a glycoside (glucinol) as the offloading substrate. This is in contrast to Tv6-931 cAT domain, which was shown to use short polyols, such as 1,1,1-tris(hydroxymethyl) ethane, pentaerythritol and glycerol, as releasing substrates.<sup>15</sup> To validate the proposed function of the LtbA-cAT domain *in vitro*, we expressed the N-terminal His-tagged cAT domain in *Escherichia coli* Rosetta(DE3) (Fig. S12†). Additionally, we synthesised 3-hydroxy-2,2,4-trimethylocta-4,6-dienoic acid *N*-acetylcysteamine thioester (3-hydroxy-SNAC) as a surrogate substrate mimicking the ACP-S-phosphopantetheinyl (Ppant)-tethered **2**. As glucinol was not readily available, we initially attempted to use its diastereomer, galactinol, as a substitute for the reaction (Fig. S7†). Various concentrations (10–50  $\mu$ M) of the recombinant stand-alone cAT domain protein were incubated with 2 mM of galactinol and **2**. However, the expected glycosylated product was not detected, possibly due to stringent substrate specificity of the LtbA-cAT domain. We next obtained authentic glucinol by hydrolysing **1** with 10 mM NaOH and repeated the reaction using the same conditions as for galactinol. However, the expected product was also not detected. This could be due to the cAT domain being unable to recognise the SNAC-tethered substrate. Thus, we attempted to assay the reverse transesterification activity of the apo-ACP-cAT domain of LtbA using an equimolar amount of **1** and free CoA or free SNAC as a mimic of the phosphopantetheinyl arm tethered on ACP (Fig. S13†). However, we could not detect the expected acyl-CoA or acyl-SNAC. This may suggest that unlike Tv6-931 PKS, which can catalyse reversible product release and recapture for double methylation, the gem-dimethyl moiety on **1** is produced by LtbA without recapture of

the monomethylated polyketide product. Hence, the ability to catalyse the reverse reaction might not be inherent to LtbA.

Next, we performed a chemical complementation assay by feeding glucinol to a culture of *A. nidulans* expressing LtbA. After 24 h, the culture was extracted with *n*-butanol and analysed by LCMS and HRESI(+)-MS (Fig. S14 and S15†). An extracted ion chromatogram (*m/z* 505) of the glucinol-fed culture revealed 7-fold enhancement of peaks corresponding to **1** and **3** as compared to a control culture that was not fed glucinol, providing further support for the hypothesis that LtbA-cAT uses glucinol as offloading substrate.

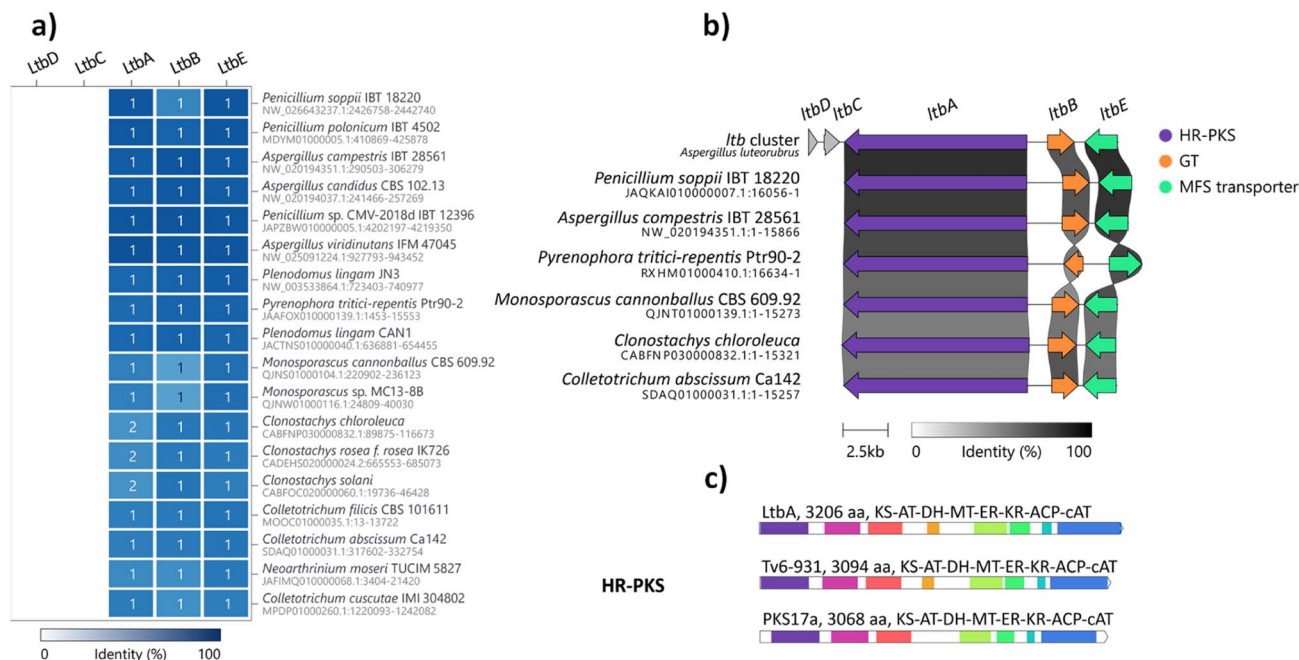
The glucinol moiety featured in **1** is unusual, hence we sought to characterise the proposed function LtbB. Amino acid sequence analysis of LtbB suggested that it is a transmembrane protein (Fig. S19 and S20†), so unsurprisingly, expression of the intron-free LtbB in *E. coli* yielded insoluble protein. Thus, cell-free lysate was prepared from *A. nidulans* expressing LtbB for the glycosylation assay. We incubated 2 mM UDP-glucose and 2 mM *myo*-inositol with the cell-free lysate in the presence of either Mg<sup>2+</sup> or Mn<sup>2+</sup>. Subsequently, the reaction was analysed using gas chromatography-mass spectrometry. However, we were not able to detect the expected product, which could be due to the lack of the correct substrate for the LtbB.

Based on all the results above, we propose a pathway for the biosynthesis of **1** (Fig. 2d). The biosynthesis starts by loading of acetyl-CoA to the ACP domain. The KS domain then catalyses the polyketide extension with malonyl-CoA loaded by the AT domain, followed by ketoreduction and dehydration by the KR and DH domains, respectively, to generate the crotonyl diketide intermediate. When the triketide intermediate is formed, a methyl group is added by the MT domain. After another elongation cycle, the tetraketide intermediate undergoes two methylations to yield the mature gem-dimethylated polyketide chain tethered on the ACP. In parallel, LtbB uses nucleoside diphosphate (NDP)-glucose as a substrate to transfer *D*-glucose to *myo*-inositol to generate glucinol. Finally, the cAT domain releases the polyketide chain from ACP, using glucinol supplied by LtbB as a nucleophile to produce **1** (Fig. S18†).

### Comparative genomic analysis reveals wide distribution of *ltb* BGC homologues in the fungal kingdom

Glycosylation is important in secondary metabolite biosynthesis, and glycosylated polyketides produced by fungi have a wide array of pharmaceutical applications. Numerous examples of fungal glycosylated polyketides have been reviewed by Hussain and co-workers.<sup>1</sup> In our study, the presence of a glucinol moiety in **1**, along with its unique release mechanism, sparked our interest in exploring homologues of the *ltb* cluster in the fungal kingdom. A search of the NCBI GenBank database using the cblaster tool<sup>22</sup> revealed homologues of the *ltb* BGC in various classes of fungi, including Eurotiomycetes, Sordariomycetes and Dothideomycetes (Fig. 3a and S21†). Selected homologous BGCs were aligned using clinker tool,<sup>23</sup> revealing uncharacterised homologues of the *ltb* BGC in several other Aspergilli, including *Aspergillus candidus*, *Aspergillus campestris* and *Aspergillus viridinutans* (Fig. 3b). Of the five genes encoded





**Fig. 3** Bioinformatic analysis of the *ltb* BGC. (a) Representative homologous BGCs found in other fungi using cblaster with high conservation ( $\geq 48\%$  identity across all genes). (b) BGCs homologous to the *ltb* BGC detected in other fungi aligned using clinker. (c) The domain architectures of the cAT domain-containing fungal HR-PKSs. LtbA in *A. luteorubrus*, Tv6-931 in *T. virens* and PKS17a in *M. anisopliae*; KS, ketosynthase; AT, malonyl-CoA acyl transferase; DH, dehydratase; MT, methyltransferase; KR, ketoreductase; ACP, acyl carrier protein; cAT, carnitine O-acyltransferase.

in the *ltb* cluster, we found three genes to be conserved in almost all of the homologous BGCs, including genes encoding the HR-PKS LtbA, the NDP-glycosyltransferase LtbB and the transporter LtbE.

Interestingly, a homologue of the *ltb* BGC appears to be conserved in the plant pathogenic fungus *Pyrenophora tritici-repentis*. As we have access to a strain of *P. tritici-repentis* (WAC13651), we selected this BGC for heterologous expression to confirm the functions of these genes. Co-expression of the LtbA HR-PKS homologue (PtrM4\_047710) and LtbB glycosyltransferase homologue (PtrM4\_047720) genes from *P. tritici-repentis* in *A. nidulans* resulted in the production of the same metabolites (*i.e.* **1**, **2** and **3**) as those encoded by the *A. luteorubrus* *ltb* BGC (Fig. S5†). This confirmed that *P. tritici-repentis*, which belongs to a different class of fungi (Dothideomycetes) to *A. luteorubrus* (Eurotiomycetes), is capable of producing luteodienosides.

### Biological activity testing of metabolites

We screened **1**, **2** and **3** for biological activity against the bacteria *Bacillus cereus* and *Staphylococcus aureus*, the fungi *Candida albicans* and *Saccharomyces cerevisiae*, and the plant *Eragrostis tef* (teff). However, no activity was observed for any of the metabolites up to  $100 \mu\text{g mL}^{-1}$  (Table S11†).

## Conclusions

In conclusion, we have discovered a novel glycosylated polyketide luteodienoside A (**1**) and uncovered the molecular basis for

its biosynthesis. Heterologous pathway reconstitution in *A. nidulans* led to **1** along with two additional products, luteodiene A (**2**) and luteodienoside B (**3**), not observed in the original fungus. In the heterologous *A. nidulans* host, the mature polyketide chain tethered on the ACP of LtbA can be hydrolysed to yield **2**, while **1** can be regioisomerised to **3**.

We have demonstrated that biosynthesis of the polyketide portion of **1** involves a novel HR-PKS LtbA capable of catalysing direct  $\alpha,\alpha$ -gem-dimethylation of the polyketide chain without requiring a product recapture step, in contrast to the previously characterised *T. virens* Tv6-931.<sup>15</sup> This is supported by phylogenetic analysis, which showed that the LtbA-cAT domain formed separate clades to other characterised cAT domains, including *T. virens* Tv6-931. Significantly, by performing PKS truncation and chemical complementation, we have provided genetic and chemical evidence that the Ltb-cAT domain utilises glucinol as a substrate to offload and decorate the polyketide chain to yield **1**. This represents the first report of a carbocyclic sugar-containing glycoside being used as an offloading nucleophile by a cAT domain of a fungal HR-PKS, thus expanding our understanding of the catalytic diversity of fungal HR-PKSs.

Additionally, we have identified an NDP-glycosyltransferase (LtbB) that is essential for the synthesis of the glucinol moiety present in **1**. Although we were unable to obtain a functional recombinant enzyme to characterise its function *in vitro*, LtbB likely serves as a glucinol synthase by transferring D-glucose to myo-inositol using NDP-glucose as a substrate. While galactinol is widespread in plants and is known to serve as an intermediary galactosyl donor for the synthesis of raffinose-containing

osmoprotectants,<sup>24</sup> there are surprisingly few studies on its diastereomer, glucinol. One such study reported the synthesis of glucinol using enzyme extracts from the basidiomycete yeast *Sporobolomyces singularis*,<sup>25</sup> while another study reported its isolation from the plants *Solanum tuberosum* and *Aegopodium podagraria*.<sup>26</sup> A series of glucinol-derived glycolipids have also been isolated from *Solanum lanceolatum*, which were shown to exhibit anti-inflammatory activities.<sup>27</sup>

Genome mining revealed that homologues of the *ltb* BGC are conserved across multiple different classes of fungi, including many plant pathogens as well as mycoparasites. We showed that the homologous *ltb* BGC in the major wheat pathogen *P. tritici-repentis* also encodes the biosynthesis of **1**. This wide distribution and the unusually high levels of production by *A. luteorubrus* (~500 mg kg<sup>-1</sup> grain) suggest an important ecological role for **1**. Given the absence of any measurable antibacterial, antifungal or herbicidal activities for **1**–**3**, this ecological role clearly warrants further investigation.

## Data availability

Additional experimental and characterisation data are available in the ESI.† Crystallographic data for compound **1** have been deposited at the Cambridge Crystallographic Data Centre (CCDC) under deposition number 2295166. The DNA sequence of the *ltb* cluster has been deposited NCBI GenBank with the accession number OR597289.

## Author contributions

A. A. A. performed the experiments and analysed the data. Z. S., H. L. and J. B. synthesised the substrate and revised the manuscript. A. C. and D. V. purified the compounds from the native strain and performed the bioassays. A. M. P. and Z. S. acquired NMR data and elucidated the structures. P. T. and W. L. obtained the crystal structure. G. R. F. acquired the experimental data from GCMS. Y. H. C., A. M. P. and E. L. supervised the study, acquired funding, and revised the manuscript. A. A. A. wrote the manuscript with input from all authors. All the authors have given approval to the final version of the manuscript.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We would like to thank A/Prof Keith Stubbs and A/Prof Ethan Goddard-Borger for helpful discussions, Prof. Berl R. Oakley (University of Kansas) for providing *A. nidulans* strain LO8030. High-resolution mass spectrometry data was acquired at the UWA Separation Science and Mass Spectrometry Platform on a Orbitrap Exploris 120 LCMS supported by Australian Research Council (ARC) (LE210100110). NMR data were acquired at the UWA Centre for Microscopy, Characterisation & Analysis (CMCA) NMR facility supported by ARC (LE190100183). We

acknowledge Genomics WA for their plasmid sequencing service. This research was funded by the ARC (DP200101880, LP210100458). The preparative HPLC system used in this study is supported by the Western Australia Future Health Research and Innovation Infrastructure Fund (FHRIFGCOVID19/14). Y.-H. C. was supported by an ARC Future Fellowship (FT160100233). A. A. A. is supported by a postgraduate scholarship from King Saud University (Saudi Arabia).

## Notes and references

- H. Hussain, N. Z. Mamadalieva, I. Ali, Elizbit, I. R. Green, D. Wang, L. Zou, J. Simal-Gandara, H. Cao and J. Xiao, *Trends Food Sci. Technol.*, 2021, **110**, 611–651.
- H. Li, C. L. M. Gilchrist, H. J. Lacey, A. Crombie, D. Vuong, J. I. Pitt, E. Lacey, Y.-H. Chooi and A. M. Piggott, *Org. Lett.*, 2019, **21**, 1287–1291.
- Y. Kasai, K. Komatsu, H. Shigemori, M. Tsuda, Y. Mikami and J. Kobayashi, *J. Nat. Prod.*, 2005, **68**, 777–779.
- A. D. Wright, C. Osterhage and G. M. König, *Org. Biomol. Chem.*, 2003, **1**, 507–510.
- F. M. Talontsi, B. Dittrich, A. Schöffler, H. Sun and H. Laatsch, *Eur. J. Org. Chem.*, 2013, **2013**, 3174–3180.
- C. L. M. Gilchrist, H. Li and Y.-H. Chooi, *Org. Biomol. Chem.*, 2018, **16**, 1620–1626.
- C.-Y. Chiang, M. Ohashi and Y. Tang, *Nat. Prod. Rep.*, 2023, **40**, 89–127.
- H. Li, A. E. Lacey, S. Shu, J. A. Kalaitzis, D. Vuong, A. Crombie, J. Hu, C. L. M. Gilchrist, E. Lacey, A. M. Piggott and Y.-H. Chooi, *Org. Biomol. Chem.*, 2021, **19**, 587–595.
- I. Roux, S. Bowles, J. A. Kalaitzis, D. Vuong, E. Lacey, Y.-H. Chooi and A. M. Piggott, *Org. Biomol. Chem.*, 2021, **19**, 9506–9513.
- H. J. Lacey, C. L. M. Gilchrist, A. Crombie, J. A. Kalaitzis, D. Vuong, P. J. Rutledge, P. Turner, J. I. Pitt, E. Lacey, Y.-H. Chooi and A. M. Piggott, *Beilstein J. Org. Chem.*, 2019, **15**, 2631–2643.
- P. W. Crous, M. J. Wingfield, Y.-H. Chooi, C. L. M. Gilchrist, E. Lacey, J. I. Pitt, F. Roets, W. J. Swart, J. F. Cano-Lira and N. Valenzuela-Lopez, *Pers.: Mol. Phylogeny Evol. Fungi.*, 2020, **44**, 301.
- J. E. Stajich and J. M. Palmer, *stajichlab/AAFTF: Bugfix Release v0.2.2*, Zenodo, 2018, DOI: [10.5281/zenodo.1658103](https://doi.org/10.5281/zenodo.1658103).
- J. M. Palmer and J. Stajich, *Funannotate v1.8.1: Eukaryotic genome annotation*, Zenodo, 2020, DOI: [10.5281/zenodo.4054262](https://doi.org/10.5281/zenodo.4054262).
- K. Blin, S. Shaw, A. M. Kloosterman, Z. Charlop-Powers, G. P. van Wezel, M. H. Medema and T. Weber, *Nucleic Acids Res.*, 2021, **49**, W29–W35.
- L. Hang, M.-C. Tang, C. J. B. Harvey, C. G. Page, J. Li, Y.-S. Hung, N. Liu, M. E. Hillenmeyer and Y. Tang, *Angew. Chem., Int. Ed.*, 2017, **56**, 9556–9560.
- Y. M. Chiang, M. Ahuja, C. E. Oakley, R. Entwistle, A. Asokan, C. Zutz, C. C. C. Wang and B. R. Oakley, *Angew. Chem., Int. Ed.*, 2016, **55**, 1662–1665.
- J. Hu, F. Sarrami, H. Li, G. Zhang, K. A. Stubbs, E. Lacey, S. G. Stewart, A. Karton, A. M. Piggott and Y.-H. Chooi, *Chem. Sci.*, 2019, **10**, 1457–1465.



- 18 I. Roux and Y.-H. Chooi, in *Engineering Natural Product Biosynthesis*, Springer, 2022, vol. 2489, pp. 75–92.
- 19 N. Sbaraini, J. Hu, I. Roux, C.-S. Phan, H. Motta, H. Rezaee, A. Schrank, Y.-H. Chooi and C. C. Staats, *Fungal Genet. Biol.*, 2021, **152**, 103568.
- 20 S. Ruswandi, K. Kitani, K. Akimitsu, T. Tsuge, T. Shiraishi and M. Yamamoto, *J. Gen. Plant Pathol.*, 2005, **71**, 107–116.
- 21 F. Kudo, Y. Matsuura, T. Hayashi, M. Fukushima and T. Eguchi, *J. Antibiot.*, 2016, **69**, 541–548.
- 22 C. L. M. Gilchrist, T. J. Booth, B. van Wersch, L. van Grieken, M. H. Medema and Y.-H. Chooi, *Bioinform. adv.*, 2021, **1**, vbab016.
- 23 C. L. M. Gilchrist and Y.-H. Chooi, *Bioinformatics*, 2021, **37**, 2473–2475.
- 24 A. Nishizawa, Y. Yabuta and S. Shigeoka, *Plant Physiol.*, 2008, **147**, 1251–1263.
- 25 P. A. J. Gorin, K. Horitsu and J. F. T. Spencer, *Can. J. Chem.*, 1965, **43**, 2259–2264.
- 26 H. Hopf and O. Kandler, *Z. Pflanzenphysiol.*, 1980, **100**, 189–195.
- 27 Y. Herrera-Salgado, M. L. Garduño-Ramírez, L. Vázquez, M. Y. Rios and L. Alvarez, *J. Nat. Prod.*, 2005, **68**, 1031–1036.

