



Cite this: *Chem. Sci.*, 2020, **11**, 11570

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

Received 5th August 2020  
Accepted 28th September 2020

DOI: 10.1039/d0sc04309e

[rsc.li/chemical-science](http://rsc.li/chemical-science)

## Introduction

Fungal maleidrides are an important family of polyketide-derived secondary metabolites which exhibit a diversity of biological activities including as antifungal<sup>1,2</sup> and herbicidal<sup>3,4</sup> agents.<sup>5</sup> They are characterised by a medium-sized alicyclic ring with one or two fused maleic anhydride moieties. The majority of the reported maleidrides are nonadrides assembled on a 9-membered ring core, and early examples include byssochlamic acid **1**,<sup>6</sup> heveadride **2**,<sup>7</sup> glauconic acid **3** and glauconic acid **4** (Fig. 1).<sup>8,9</sup> Later studies revealed further nonadrides such as the phomoidrides, *e.g.* **5**,<sup>10</sup> castaneolide **6**,<sup>11</sup> rubratoxins, *e.g.* **7**<sup>12</sup> and cornexistin **8**,<sup>3,13</sup> the octadride viburspiran **9** and more recently two heptadrides agnestadrides A **10** and B have been isolated from *Byssochlamys fulva*.<sup>14</sup> In certain cases the relative

<sup>a</sup>School of Chemistry, University of Bristol, Cantock's Close, Bristol, BS8 1TS, UK.  
E-mail: [Chris.Willis@bristol.ac.uk](mailto:Chris.Willis@bristol.ac.uk); [andy.bailey@bristol.ac.uk](mailto:andy.bailey@bristol.ac.uk); [kd4495@bristol.ac.uk](mailto:kd4495@bristol.ac.uk)

<sup>b</sup>School of Biological Sciences, University of Bristol, 24 Tyndall Avenue, Bristol, BS1 1TQ, UK

<sup>c</sup>Syngenta, Jealott's Hill International Research Centre, Bracknell, RG42 6EY, UK

<sup>d</sup>Institute for Organic Chemistry and BMWZ, Leibniz University of Hannover, Schneiderberg 38, 30167, Hannover, Germany

<sup>†</sup> Electronic supplementary information (ESI) available. CCDC 2011262–2011264. Crystallographic data for compounds **16**, **17** and **18**. GenBank accession numbers; MT724049 (zopfiellin BGC) and MT724050 (scytalidin BGC). For ESI and crystallographic data in CIF or other electronic format see DOI: 10.1039/d0sc04309e

<sup>‡</sup> These authors contributed equally.

and absolute stereochemistry of the maleidrides remain unknown, and various details of their biosynthesis are yet to be fully elucidated.

Results of extensive studies on selected nonadrides using isotopic labelling, gene knock outs and heterologous expression experiments have led to the proposal that the biosynthetic pathway begins with the assembly of an unsaturated precursor **11**

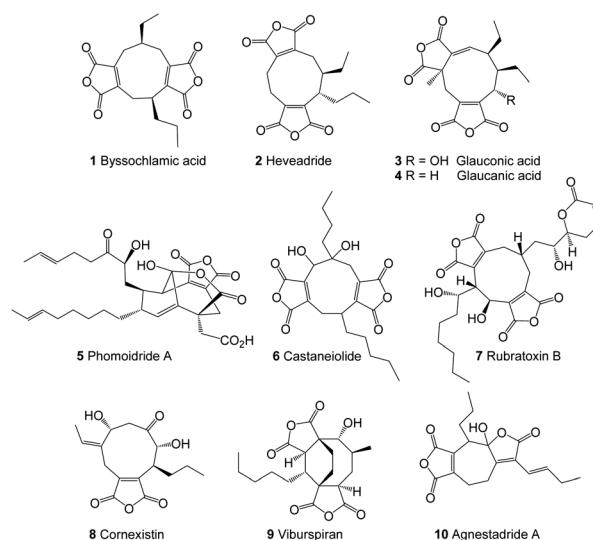
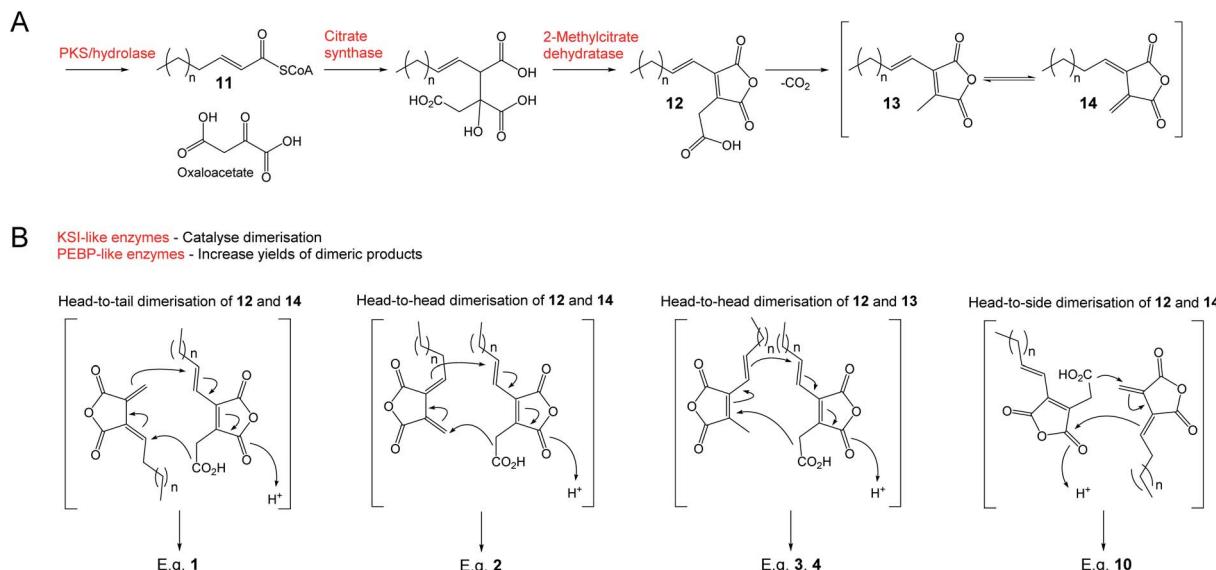
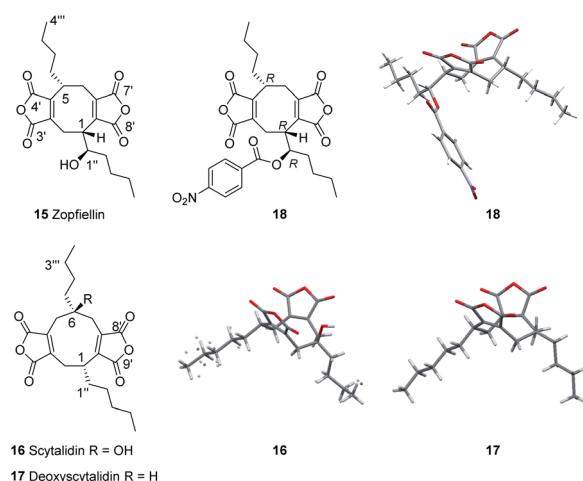


Fig. 1 Examples of maleidrides.

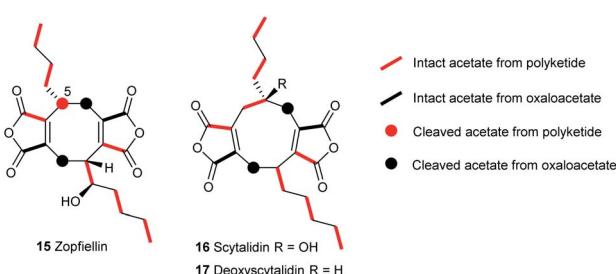




**Scheme 1** The universal pathway to maleidrides<sup>16</sup> involves the production of maleic anhydride monomers (A) followed by dimerisation (B) – the mode of which determines the carbon framework of the final compound.



**Fig. 2** Confirmed structures of the natural products 15, 16 and 17.



**Fig. 3** Incorporation patterns in octadrides and nonadrides from feeding studies with  $[1,2-^{13}\text{C}_2]$ -acetate.

via an iterative highly reducing polyketide synthase (hrPKS), where the chain length varies according to the structure of the natural product (Scheme 1). Coupling of 11 with oxaloacetate is

catalysed by citrate synthase-like enzymes and is followed by dehydration catalysed by 2-methylcitrate dehydratases to generate the maleidride monomer 12.<sup>15,16</sup> Decarboxylation of 12 then gives the tautomeric compounds 13 and 14 (Scheme 1A).<sup>14,17</sup> The second stage of the biosynthetic pathway, coupling the monomers in various modes and cyclisation to the carbocyclic ring, involves ketosteroid isomerase-like (KSI) and phosphatidylethanolamine binding protein-like (PEBP) enzymes.<sup>16</sup> The mode of cyclisation determines the carbon framework of the maleidride (Scheme 1B), and finally tailoring modifications decorate these structures leading to the observed diversity of natural products.

Octadrides, with 8-membered rings, are less common than nonadrides and include viburspiran 9<sup>18</sup> and zopfiellin 15<sup>19</sup> (Fig. 1 and 2). The biosynthetic origin of the 8-membered ring is unknown (§see footnote). Zopfiellin was first isolated from *Zopfiella curvata* and the structure determined by a combination of spectroscopic methods.<sup>19-21</sup> The relative and absolute configurations of zopfiellin have been indicated in several publications,<sup>22-24</sup> but we were unable to find any justification of the configuration of the side chain secondary alcohol. Hence, with our continuing interest in the biosynthesis of fungal natural products, and in particular bioactive maleidrides, our goal was to confirm the structure of zopfiellin 15 and determine the biosynthetic origin of the 8-membered ring, with the longer-term goal of engineering the pathway to new antifungal targets.

Whilst this manuscript was in preparation Minami, Oikawa and co-workers reported investigations into the biosynthesis of zopfiellin using isotopic labelling, heterologous expression of candidate genes in *Aspergillus oryzae*, combined with structure elucidation of minor metabolites. These studies showed that a ring contraction occurs from a nonadride to an octadride.<sup>25</sup> Herein we describe our recent studies which confirm the absolute and relative configurations of the nonadride scytalidin 16 and octadride zopfiellin 15 and reveal that they share

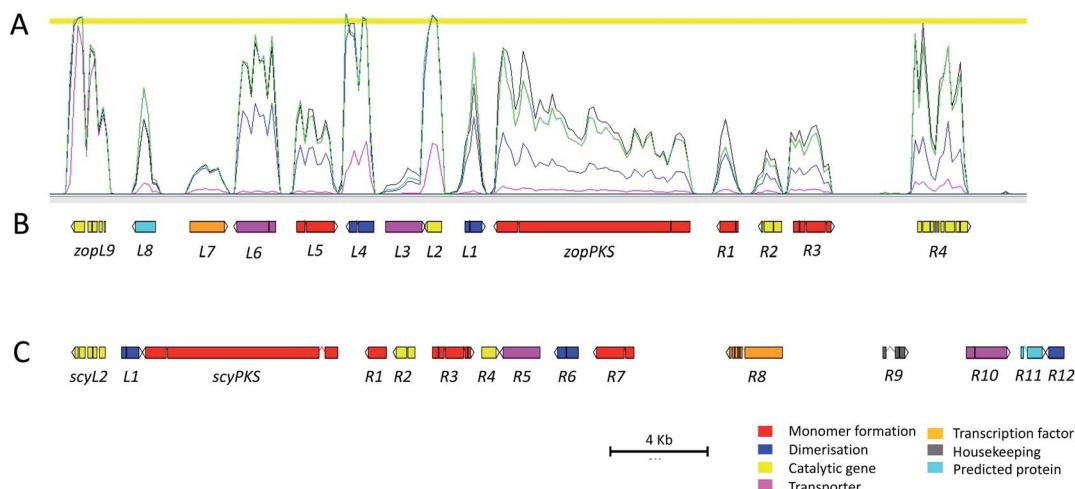


Fig. 4 (A) RNAseq data mapped to the zopfiellin BGC showing coregulation of the genes under four different conditions (Fig. S1 and S2†). Pink: GN. Purple: CDB. Green: PDB-day 5. Black: PDB-day 8. (B) The zopfiellin gene cluster, as annotated in Artemis, which contains genes *zopL9* to *zopR4*. (C) The scytalidin gene cluster, as annotated in Artemis, which contains genes *scyL2* to *scyR12*. 'L' in the gene names denotes genes to the left, or upstream, of the PKS. 'R' denotes genes to the right, or downstream, of the PKS. Putative gene functions are shown in Table 1.

a common biosynthetic precursor, deoxyscytalidin 17 (Fig. 2). Our biosynthetic investigations involved a different approach from those previously reported<sup>25</sup> and used a combination of gene disruptions in *Diffractella curvata* and *Scytalidium album*, feeding studies to mutant strains and full structure elucidation of known and novel compounds. These studies establish the biosynthetic relationships between zopfiellin 15, scytalidin 16, and deoxyscytalidin 17.

## Results and discussion

The confirmed zopfiellin producer, *Zopfiella curvata* no. 37-3, was not publicly available when these studies began, so

*Diffractella curvata* CBS591.74 was obtained and confirmed to produce zopfiellin. Cultures were grown in shake flasks in PDB medium at 25 °C for 14 days and after purification by flash chromatography, gave zopfiellin 15 (80–100 mg L<sup>-1</sup>) as an oil. As full assignment of NMR data had not previously been reported<sup>19,21</sup> 2D NMR (COSY, HSQC and HMBC) analysis allowed assignment of the <sup>1</sup>H and <sup>13</sup>C NMR spectra (Table S6 and Fig. S31–S36†). The optical rotation, [α]<sub>D</sub> −84.3 (c 0.43 MeOH) was consistent with the literature value, [α]<sub>D</sub> −76.8 (c 0.42 MeOH).<sup>20</sup> Zopfiellin was esterified to give the novel *p*-nitrobenzoate 18 and X-ray crystallography revealed the relative configuration of the natural product (Fig. 2). Synthesis and analysis of the (R)- and (S)-Mosher's ester derivatives of 15

Table 1 Predicted functions for genes in the zopfiellin and scytalidin BGCs. Function were determined by searching for homologues within the Swissprot database and identifying conserved protein domains using InterPro.<sup>33</sup> Only the functions of genes marked '\*' were assigned via the identification homologues within NCBI's non-redundant database. See Tables S1 and S2 for further details

Gene	Putative function	Homologue	Gene	Putative function	Homologue
<i>zopL9</i>	Hydroxylase/Desaturase	DES <sup>34</sup> AsaB <sup>35</sup>	<i>scyL2</i>	Hydroxylase	AclN <sup>36</sup> MfR1 <sup>37</sup>
<i>zopL8</i>	Hypothetical protein	—	<i>scyL1*</i>	PEBP	BfL5 <sup>16</sup>
<i>zopL7</i>	Transcription factor (TF)	AlnR <sup>38</sup>	<i>scyPKS</i>	hrPKS	Tox1-PKS <sup>39</sup>
<i>zopL6</i>	Major facilitator superfamily (MFS) transporter	Itp1 <sup>40</sup>	<i>scyR1</i>	DUF341 hydrolase/esterase	Fub4 <sup>41</sup>
<i>zopL5</i>	2-Methylcitrate dehydratase (2MCDH)	PrpD <sup>42</sup>	<i>scyR2</i>	Enoyl CoA hydratase	Ech1 <sup>43</sup>
<i>zopL4*</i>	Ketosteroid isomerase-like protein (KSI)	BfL6 <sup>16</sup>	<i>scyR3</i>	Citrate synthase	MfR3 <sup>44</sup>
<i>zopL3</i>	Major facilitator superfamily (MFS) transporter	Itp1 <sup>40</sup>	<i>scyR4</i>	Isochorismatase-like hydrolase/amidohydrolase	CSHase <sup>45</sup>
<i>zopL2</i>	Isochorismatase-like hydrolase/amidohydrolase	NicF <sup>46</sup>	<i>scyR5</i>	MFS transporter	MfM6 <sup>44</sup>
<i>zopL1</i>	Phosphatidylethanolamine-binding protein (PEBP)	Tfl1 <sup>47</sup>	<i>scyR6*</i>	KSI	BfL6 <sup>16</sup>
<i>zopPKS</i>	hrPKS	Tox1-PKS <sup>39</sup>	<i>scyR7</i>	2-Methylcitrate dehydratase	PrpD <sup>42</sup>
<i>zopR1</i>	DUF341 hydrolase/esterase	Fub4 <sup>41</sup>	<i>scyR8</i>	Transcription factor	AlnR <sup>38</sup>
<i>zopR2</i>	Enoyl-CoA hydratase	Ech1 <sup>43</sup>	<i>scyR9</i>	Histone H2A	H2A.Z <sup>48</sup>
<i>zopR3</i>	Citrate synthase	CshA <sup>49</sup>	<i>scyR10</i>	MFS transporter	MfR5 <sup>44</sup>
<i>zopR4</i>	FAD-dependent oxidoreductase	YanF <sup>50</sup> Sols <sup>51,52</sup>	<i>scyR11</i>	Hypothetical protein	—
			<i>scyR12*</i>	PEBP	BfL9 <sup>16</sup>



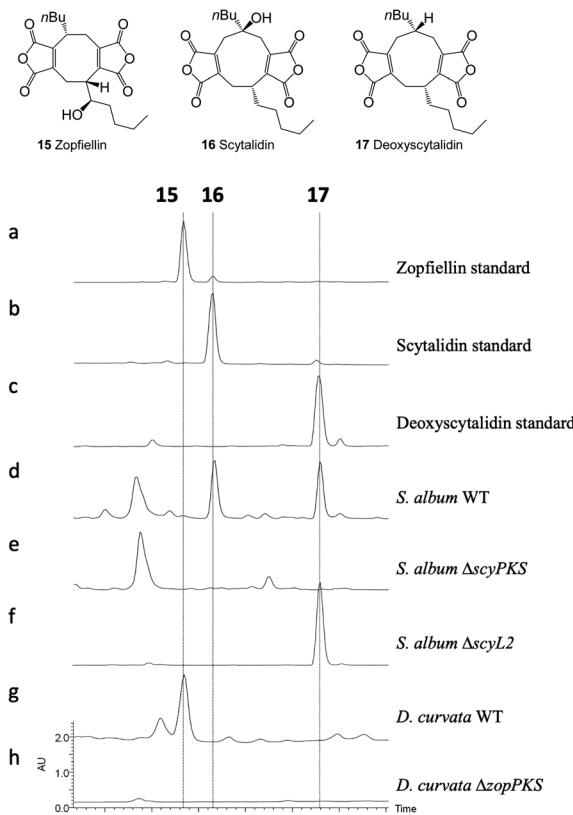


Fig. 5 HPLC (DAD) analysis of gene deletion strains for the zopfiellin and scytalidin gene clusters. Disruption of *zopPKS* and *scyPKS* abolished maleidride production in both species, confirming the identity of the zopfiellin and scytalidin BGCs. Disruption of *scyL2* confirmed its role in the final hydroxylation of the scytalidin pathway.

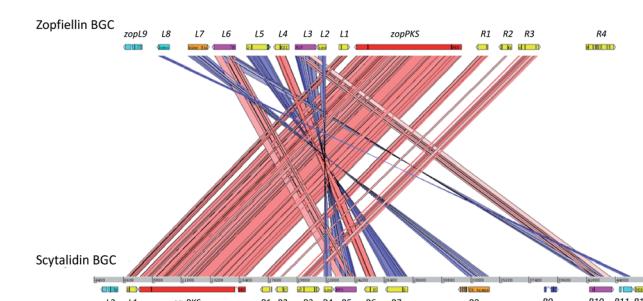


Fig. 6 An ACT (Artemis Comparison Tool) comparison of the zopfiellin and scytalidin BGCs identified genes predicted to encode homologous proteins and highlighted the similarity between the two biosynthetic gene clusters.

confirmed the absolute stereochemistry of zopfiellin as 1*R*, 5*R*, 1''*R* (Fig. S81–S83†).

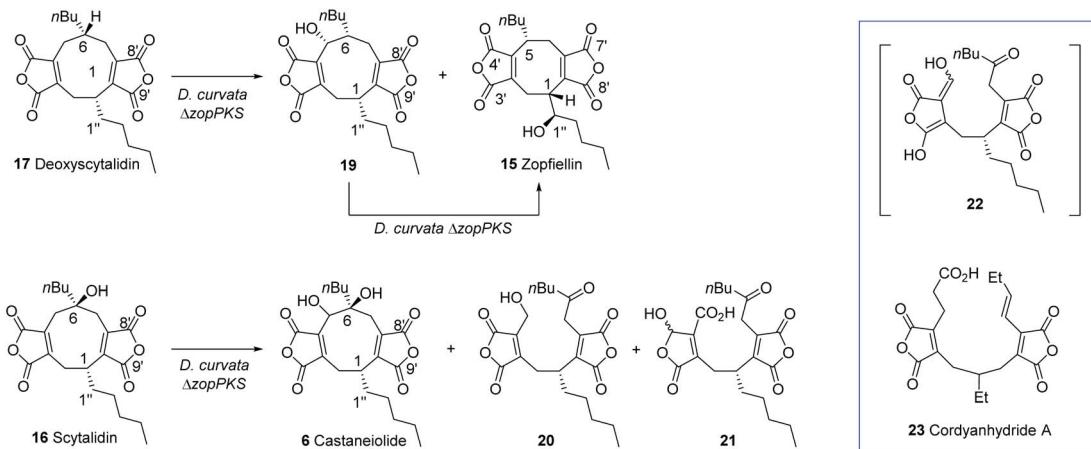
Feeding [1,2-<sup>13</sup>C<sub>2</sub>]-acetate to cultures of *D. curvata* showed mainly intact incorporation of acetate into zopfiellin 15 (Fig. 3). Enhancement of C-5 in the <sup>13</sup>C NMR spectrum was observed, but there was no coupling to an adjacent <sup>13</sup>C, indicating that it originates from a cleaved acetate and therefore that a carbon atom is lost during zopfiellin biosynthesis.

A possible biosynthetic precursor to zopfiellin is therefore a nonadride which on rearrangement with the loss of one carbon atom would give the corresponding octadride. Such ring contractions are well preceded in fungal biosynthesis, for example, in the contraction of ring B of the kaurenoids to the gibberellins,<sup>26</sup> the conversion of the 14-membered polyketide macrodiolide colletotriene to the 13-membered ring of bartanol<sup>27</sup> and the oxidative ring contractions seen in xenovulene biosynthesis.<sup>28</sup> Conceivably the known natural products scytalidin 16<sup>29</sup> (C<sub>22</sub>H<sub>28</sub>O<sub>7</sub>), deoxyscytalidin 17<sup>30</sup> (C<sub>22</sub>H<sub>28</sub>O<sub>6</sub>) or castaneiolide 6<sup>11</sup> (C<sub>22</sub>H<sub>28</sub>O<sub>8</sub>) could be precursors. To obtain samples of 16 and 17, cultures of *S. album* strain UAMH 3620 were grown, and after 14 days compounds 16 and 17 were isolated in similar titres (ca. 10 mg L<sup>-1</sup>). X-ray analysis confirmed that the alkyl side chains are *syn* in both metabolites (Fig. 2). The results of feeding studies using [1,2-<sup>13</sup>C<sub>2</sub>] acetate to cultures of *S. album* were in accord with the predicted biosynthetic pathway to nonadrides and that one or both could possibly be biosynthetic precursors of zopfiellin 15 (Fig. 3).

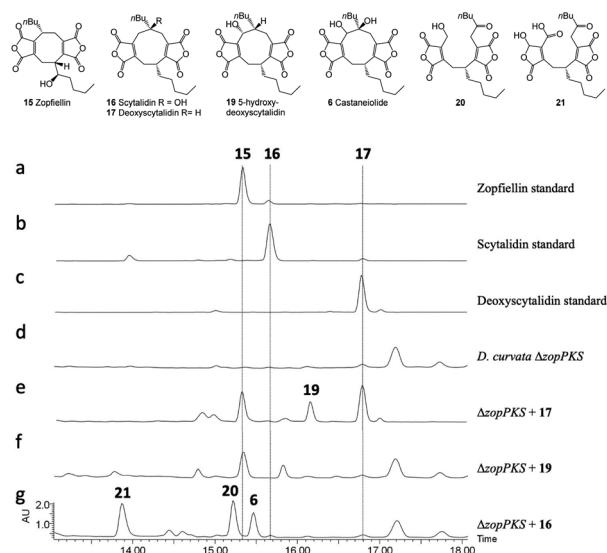
*De novo* genome sequencing was performed for *D. curvata* CBS591.74 and *S. album* UAMH3620, and an initial screen of the genome data, searching for homologues to the byssochlamic acid biosynthetic gene cluster (BGC),<sup>16</sup> revealed putative maleidride gene clusters within both genomes. The gene clusters were analysed in detail using FGeneSH<sup>31</sup> to predict coding sequences, and manually annotated in Artemis<sup>32</sup> to produce full draft BGCs (Fig. 4 and Table 1). Putative gene functions were assigned via the identification of homologues in the NCBI database and the detection of conserved domains using InterPro<sup>33</sup> (Tables S1 and S2†). Transcriptomic analysis for *D. curvata* under a range of zopfiellin 15 production and non-production conditions aided identification of the cluster boundaries by clearly identifying a co-transcribed region that includes genes from *zopL9* to *zopR4* (Fig. 4, S1 and S2†). In addition, detailed analysis of the transcriptomic data confirmed the positions of all introns and exons in the predicted ORFs. However, despite repeated attempts, it was not possible to generate quality RNAseq data for the putative scytalidin gene cluster in *S. album*.

To confirm the identity of the two gene clusters, protoplast-mediated transformation protocols were developed for both fungal species and the PKS genes; *zopPKS* and *scyPKS*, were disrupted using the bipartite gene knock-out approach developed by Nielsen *et al.*<sup>53</sup> In both cases this led to a total loss of maleidride biosynthesis, with no scytalidin 16 or deoxyscytalidin 17 being present in the crude extracts of the *ΔscyPKS* strains (Fig. 5e), and no zopfiellin 15 being present in the crude extracts of the *ΔzopPKS* strains (Fig. 5h).

An ACT (Artemis Comparison Tool)<sup>54</sup> comparison of the zopfiellin and scytalidin BGCs (Fig. 6) highlighted significant homologies. Both clusters contain genes encoding the essential hrPKS, DUF341 hydrolase, citrate synthase (CS) and 2-methylcitrate dehydratase (2MCDH),<sup>12,13,15,16</sup> as well as the 'dimerisation' genes encoding a ketosteroid-isomerase-like protein (KSI) and a PEBP-like protein.<sup>16</sup> In the case of the scytalidin BGC a second PEBP gene is present, which has been observed previously in the byssochlamic acid and rubratoxin BGCs.<sup>12,16</sup> A domain analysis of the two PKS genes identified all of the



**Scheme 2** Feeding studies to *D. curvata* strain  $\Delta zopPKS$  identified **17** and **19** as intermediates in zopfiellin biosynthesis and scytalidin was metabolised to novel products.



**Fig. 7** HPLC (DAD) chromatograms of crude extracts. Feeding compounds to *D. curvata*  $\Delta zopPKS$  indicate that deoxyscytalidin **17** and 5-hydroxy-deoxyscytalidin **19** are intermediates in the zopfiellin biosynthetic pathway, whereas scytalidin **16** is not. Feeding of **16** led to the production of castaneolide **6** and two novel ring-open anhydrides **20** and **21**.

domains typically present in a highly-reducing PKS: KS, AT, DH, CMeT, ER, KR and ACP. Additional genes identified in both clusters encode an enoyl-CoA hydratase, an isochorismatase, a small hypothetical protein (*zopL8* and *scyR11*), a GAL4 type transcription factor and two MFS (major facilitator superfamily) transporters. The enoyl-CoA hydratase encoded by the zopfiellin cluster (*zopR2*) was not identified in the work by Oikawa and colleagues,<sup>25</sup> but RNAseq data confirmed its annotation.

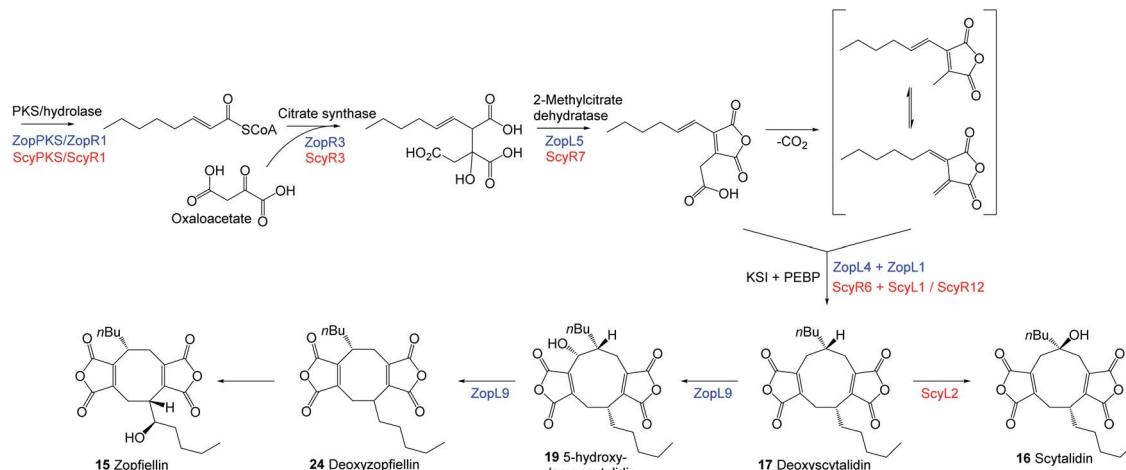
The proteins encoded by the clusters also have high sequence identity (Table S3†). The KSI and PEBP-like proteins, for example, are far more similar to one another (80.7% and 66.2% identity respectively, Table S3†) than to those identified in other maleidride BGCs (Fig. S4–S7†). In the case of the

transporters, ZopL6 and ScyR5 are highly homologous (80.4% identity) and thus may be expected to transport similar compounds. The other two transporters, ZopL3 and ScyR10, are more diverse, sharing only 38.5% identity. All other homologues encoded by the gene clusters share over 50% identity, with many sharing over 80% (Table S3†).

The only proteins which are uniquely encoded by the zopfiellin gene cluster are a putative hydroxylase/desaturase (ZopL9), which shares homology with DES from the gibberellin pathway, and an FAD-dependent oxidoreductase (ZopR4). The scytalidin gene cluster encodes only one unique protein, a putative hydroxylase (ScyL2). ZopL9 and ScyL2 are both predicted to be  $\alpha$ -ketoglutarate ( $\alpha$ KG) dependent enzymes based on the presence of conserved iron and  $\alpha$ KG binding residues (Fig. S10†), but they were not considered to be homologues due to the low level of sequence identity (approximately 25% identity), suggesting that the exact catalytic function is unlikely to be the same. Also located within the scytalidin BGC is a gene (*scyR9*) encoding a highly conserved H2A.F/Z family histone,<sup>55</sup> but this is unlikely to play a direct role in scytalidin biosynthesis (Table S3†).

The striking similarity between the two gene clusters (Fig. 6) suggests that zopfiellin **15** and scytalidin **16** have very similar biosynthetic origins, and that either deoxyscytalidin **17** or scytalidin **16** may be precursors of zopfiellin. To test this theory, **16** and **17** were purified from *S. album* cultures and separately fed to *D. curvata*  $\Delta zopPKS$  cultures on days 3 and 5 post-inoculation (5 mg per day per 100 ml of culture). In cultures fed with **17**, zopfiellin production was restored (Scheme 2 and Fig. 7e), demonstrating that deoxyscytalidin **17** is a precursor of zopfiellin. The extent of conversion of deoxyscytalidin to zopfiellin varied between individual feeds, presumably due to varying expression levels for the zopfiellin BGC. In cultures where the conversion was partial, the HPLC trace showed a peak for deoxyscytalidin, as well as a novel compound with a mass of 404 eluting at 16.1 minutes (Fig. 7e). This metabolite was isolated, and the structure confirmed by NMR spectroscopy to be 5-hydroxy-deoxyscytalidin **19** (Table S9 and Fig. S44–S49†). The





Scheme 3 Proposed biosynthetic pathways of the maleidrides deoxyscytalidin 17, scytalidin 16 and zopfiellin 15.

signal assigned to 5-H (at  $\delta$  4.90) appeared as a broad singlet and NOE studies were in accord with hydroxylation occurring on the same face as the butyl side chain at C-6. To determine whether **19** is a shunt product or an intermediate on the biosynthetic pathway, it was fed to *D. curvata*  $\Delta$ zopPKS cultures and again this restored zopfiellin biosynthesis (Scheme 2 and Fig. 7f).

When scytalidin **16** was fed to cultures of *D. curvata*  $\Delta$ zopPKS, no zopfiellin was detected, but several additional compounds were present when compared to controls (Fig. 7g). The mixture was purified by HPLC giving the known<sup>11</sup> nonadride castaneolide **6** as one of the metabolites, presumably formed by 5-hydroxylation of scytalidin **16** (Scheme 2). In addition, two novel compounds **20** and **21** were isolated (3.2 mg and 0.7 mg respectively) and their structures elucidated by MS and extensive NMR studies (Table S10 and Fig. S50–S63†). The  $^{13}\text{C}$  NMR spectrum of compound **20** ( $\text{C}_{22}\text{H}_{28}\text{O}_8$  from HRMS) showed downfield signals assigned to the two maleic anhydrides, as well as an additional signal at  $\delta_{\text{C}}$  205 ppm assigned to a ketone. The  $^1\text{H}$ -NMR showed characteristic AB signals assigned to the hydroxymethylene. Based on this evidence, and combined with further analysis of 2D NMR data, the new metabolite was assigned as the ring cleaved alcohol **20**. The spectral data of **21** was also in accord with a ring cleaved metabolite with a ketone in the side chain, but in this case the primary alcohol was oxidised to a carboxylic acid and one of the carbonyls of the anhydride reduced to give a mixture of epimers ( $\delta_{\text{C}}$  96.9/97.4 ppm,  $\delta_{\text{H}}$  6.19/6.23 ppm). Such reductions of anhydrides are commonly observed in maleidride biosynthesis.<sup>14,56,57</sup> We propose that ketone **20** may be formed *via* cleavage of the 5,6-diol **6** to give an intermediate **22**, which would tautomerise to generate ketone **20**. Oxidation of the primary alcohol of **20** to a carboxylic acid and selective reduction of the anhydride would give **21**. Such linear bis(maleic anhydride) structures are not without precedent, for example the fungal metabolite cordyanhydride A **23**.<sup>58</sup> However, **23** has previously been proposed to arise *via* the linear coupling of two distinct maleic anhydride monomers rather than *via* cleavage of a cyclic intermediate.<sup>15</sup>

Based on the bioinformatic analysis of the BGCs and the identification of deoxyscytalidin **17** as an intermediate to both scytalidin **16** and zopfiellin **15**, it is a reasonable inference that the two biosynthetic pathways proceed in an identical manner to **17**, after which they diverge. Hydroxylation at C-6 would produce scytalidin whereas hydroxylation at C-5, ring contraction and hydroxylation of the side chain would produce zopfiellin.

ScyL2, being the only unique catalytic protein encoded by the scytalidin BGC, was identified as a likely candidate for catalysing the final step in the scytalidin biosynthetic pathway. This is supported by a bioinformatic analysis identifying ScyL2 as a likely  $\alpha$ -ketoglutarate dependent hydroxylase (Table S2†), which shares homology with enzymes known to catalyse hydroxylations, such as MfR1 and MfR2 from the squalenol pathway,<sup>37</sup> and RbtB and RbtG from the rubratoxin pathway<sup>12</sup> (Fig. S8–S10†). To confirm this role, scyL2 knock-out strains of *S. album* were generated. These strains did not produce scytalidin **16**, but accumulated deoxyscytalidin **17** (Fig. 5f), confirming the role of ScyL2.

Recent work by Shiina *et al.*<sup>25</sup> on the zopfiellin pathway of *Z. curvata*, using heterologous expression as well as *in vitro* assays, has begun to elucidate the later stages of zopfiellin biosynthesis. ZopK, which is equivalent to ZopL9 of the *D. curvata* BGC reported here, was shown to catalyse the oxidative ring contraction in the zopfiellin pathway to produce deoxyzopfiellin **24**.<sup>25</sup> This proceeds *via* the intermediate **19** (which we now know to be 5-hydroxy-deoxyscytalidin) and is consistent with our identification, *via* feeding studies, of **19** as an intermediate in the zopfiellin pathway. *In vitro* assays using purified recombinant ZopL9 gave results which are entirely consistent with the findings of Shiina *et al.*<sup>25</sup> In the presence of  $\alpha$ -ketoglutarate, ZopL9 was capable of catalysing highly efficient hydroxylation of deoxyscytalidin **17** to give 5-hydroxy-deoxyscytalidin **19** (Fig. S68†). A minor compound was also detected which was purified and identified as deoxyzopfiellin **24**. Repeating the assays using 5-hydroxy-deoxyscytalidin **19** as the substrate also gave rise to deoxyzopfiellin **24** (Fig. S68†), in accord with the

proposal that ZopL9 is implicated in the ring contraction in the zopfiellin pathway of *D. curvata*, by first hydroxylating deoxyscytalidin to give **19**, then catalysing the ring contraction to give the octadride core.

Having identified **17**, **19** and **24** as intermediates in the zopfiellin pathway – through both feeding studies and *in vitro* assays – an extract of wild-type *D. curvata* was re-analysed, which revealed the presence of all three compounds as minor metabolites (Fig. S25†). Taken together with the bioinformatic analysis and the discovery that ScyL2 catalyses the final step of the scytalidin pathway, a branching biosynthetic route for the production of deoxyscytalidin, scytalidin and zopfiellin can now be proposed as shown in Scheme 3.

Intriguingly there is a homologue to ZopL9 (PhiK) encoded by the phomoidride BGC (44.2% homology – Fig. S11†). The structure of the phomoidrides, *e.g.* **5**, suggests that multiple oxidations must occur during their biosynthesis, but no ring contraction of the nine-membered core occurs in this pathway, and the role of PhiK has not yet been investigated.

## Conclusions

This work categorically establishes the biosynthetic relationship between the nonadride scytalidin **16** and the octadride zopfiellin **15**, *via* the shared nonadride intermediate; deoxyscytalidin **17**. In both pathways, deoxyscytalidin is the substrate for  $\alpha$ KG-dependent dioxygenase enzymes. In the scytalidin pathway ScyL2 catalyses a 6-hydroxylation of deoxyscytalidin to produce scytalidin whereas in the zopfiellin pathway ZopL9 catalyses a 5-hydroxylation of deoxyscytalidin and is implicated in a ring contraction to produce the octadride core. Interestingly, when cultures of *D. curvata*  $\Delta$ zopPKS were fed with scytalidin, (possessing a 6-hydroxy group), 5-hydroxylation occurred to give castaneolide **6** and two novel ring-cleaved maleidrides **20** and **21** were produced. Genes encoding  $\alpha$ KG-dependent enzymes have been identified within a further three maleidride BGCs; namely *rbtB*, *rbtE*, *rbtG* and *rbtU* of the rubratoxin cluster, and the currently uncharacterised *phiK* of the phomoidride BGC and *pvl5* of the cornexistin BGC. This highlights the important role these, often multifunctional, enzymes play in generating the structural diversity seen in the maleidride class of natural products. Intriguing questions which remain to be answered regarding the biosynthesis of zopfiellin include the mechanism by which the ring contraction occurs, and the identification of the enzyme responsible for installing the hydroxyl group at C-1" of the side chain of zopfiellin. ZopR4, which is uniquely encoded by the zopfiellin BGC and predicted to be a FAD-dependent oxidoreductase, is a candidate for catalysing such a transformation, but this remains to be explored experimentally. Additionally, the roles of the isochorismatase-like enzyme (ZopL2/ScyR4), the enoyl CoA hydratase (ZopR2/ScyR2) and the small hypothetical protein (ZopL8/ScyR11), which are encoded by both BGCs, are currently unknown.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We thank BBSRC (BB/K002341/1) and Syngenta for funding (KMJdMS, CG and ZS) and a PhD studentship BB/P504804/1 (CS). We are also grateful to MRC (MR/N029909/1) (funding for TTD) and EPSRC (EP/L015366/1), Bristol Chemical Synthesis Centre for Doctoral Training, which provided PhD studentships (DMH and DOF). We are also very grateful to BBSRC and EPSRC for funding instrumentation *via* the Bristol Centre for Synthetic Biology (BB/L01386X/1). Genome sequencing, the production of assembled draft genomes and RNA sequencing were all carried out at the DNA sequencing facility in the Biochemistry Department of Cambridge University. X-ray crystallography was carried out by Dr Hazel Sparkes, School of Chemistry, University of Bristol. In addition we are grateful to Dr Lauren Ray from Syngenta for her help and support during the Industrial placement for CS.

## Notes and references

§ The numbering systems used for the maleidrides varies greatly in the literature and shows no consistency. We propose a common more systematic system based on the size of the ring (1–9, 1–8, 1–7 as appropriate) beginning at the carbon alpha to the maleic anhydride ring, which gives the lowest numbers to the side chains. The maleic anhydride carbons would be numbered with a prime, appropriate to the ring numbering, hence 3'4' and 8'9' for byssochlamic acid, and 1", 2", *etc.* for the first side chain, numbering from the ring junction and 1", 2", *etc.* for the second side chain.

- 1 T. Hosoe, K. Fukushima, T. Itabashi, K. Nozawa, K. Takizawa and K.-I. Kawai, *Heterocycles*, 2004, **63**, 2581–2589.
- 2 M. Futagawa, A. M. Rimando, M. R. Tellez and D. E. Wedge, *J. Agric. Food Chem.*, 2002, **50**, 7007–7012.
- 3 M. Nakajima, K. Itoi, Y. Takamatsu, S. Sato, Y. Furukawa, K. Furuya, T. Honma, J. Kadotani, M. Kozasa and T. Haneishi, *J. Antibiot.*, 1991, **44**, 1065–1072.
- 4 S. C. Fields, L. Mireles-Lo and B. C. Gerwick, *J. Nat. Prod.*, 1996, **59**, 698–700.
- 5 R. Schor and R. Cox, *Nat. Prod. Rep.*, 2018, **35**, 230–256.
- 6 H. Raistrick and G. Smith, *Biochem. J.*, 1933, **27**, 1814–1819.
- 7 R. I. Crane, P. Hedden, J. MacMillan and W. B. Turner, *J. Chem. Soc., Perkin Trans. 1*, 1973, 194–200.
- 8 N. Wijkman, *Justus Liebigs Ann. Chem.*, 1931, **485**, 61–73.
- 9 J. E. Baldwin, D. H. R. Barton, J. L. Bloomer, L. M. Jackman, L. Rodriguez-Hahn and J. K. Sutherland, *Experientia*, 1962, **18**, 345–352.
- 10 P. Spencer, F. Agnelli and G. A. Sulikowski, *Org. Lett.*, 2001, **3**, 1443–1445.
- 11 K. Arai, S. Shimizu, H. Miyajima and Y. Yamamoto, *Chem. Pharm. Bull.*, 1989, **37**, 2870–2872.
- 12 J. Bai, D. Yan, T. Zhang, Y. Guo, Y. Liu, Y. Zou, M. Tang, B. Liu, Q. Wu, S. Yu, Y. Tang and Y. Hu, *Angew. Chem., Int. Ed.*, 2017, **56**, 4782–4786.
- 13 K. Williams, A. J. Szwabe, C. Dickson, T. R. Desson, N. P. Mulholland, J. L. Vincent, J. M. Clough, A. M. Bailey, C. P. Butts and C. L. Willis, *Chem. Commun.*, 2017, **53**, 7965–7968.



14 A. J. Szwabe, K. Williams, D. E. O'Flynn, A. M. Bailey, N. P. Mulholland, J. L. Vincent, C. L. Willis, R. J. Cox and T. J. Simpson, *Chem. Commun.*, 2015, **51**, 17088–17091.

15 R. Fujii, Y. Matsu, A. Minami, S. Nagamine, I. Takeuchi, K. Gomi and H. Oikawa, *Org. Lett.*, 2015, **17**, 5658–5661.

16 K. Williams, A. J. Szwabe, N. P. Mulholland, J. L. Vincent, A. M. Bailey, C. L. Willis, T. J. Simpson and R. J. Cox, *Angew. Chem., Int. Ed.*, 2016, **55**, 6784–6788.

17 D. M. Heard, E. R. Tayler, R. J. Cox, T. J. Simpson and C. L. Willis, *Tetrahedron*, 2020, **76**, 130717.

18 M. Saleem, H. Hussain, I. Ahmed, S. Draeger, B. Schulz, K. Meier, M. Steinert, G. Pescitelli, T. Kurtan, U. Flörke and K. Krohn, *Eur. J. Org. Chem.*, 2011, **2011**, 808–812.

19 Y. Nakajima, H. Watanabe, T. Adachi, M. Tagawa, M. Futagawa and Y. Nishino, Japan Patent, JP08107792, 1996.

20 T. Watanabe, T. Yasumoto, M. Murata, M. Tagawa, H. Narushima, T. Furusato, M. Kuwahara, M. Hanaue and T. Seki, Europe Patent, EP0582267A1, 1994.

21 A. Nagamura, H. Aoki, T. Nishimura and I. Kinoshita, Japan Patent, JP09087268, 1997.

22 Y. Baba, C. Noutary, S. Ichikawa, T. Kusumoto, H. Watanabe, M. Adachi, H. Hotta, Y. Nakajima, K. Arai and T. Hiyama, *Synlett*, 1997, **12**, 1393–1394.

23 H. Watanabe, Y. Nakajima, M. Adachi, H. Hotta, K. Arai, Y. Baba, C. Noutary, S. Ichikawa and T. Kusumoto, *Chem. Commun.*, 1999, **17**, 1753–1754.

24 L. Musso, S. Dallavalle, G. Farina and E. Burrone, *Chem. Biol. Drug Des.*, 2012, **79**, 780–789.

25 T. Shiina, T. Ozaki, Y. Matsu, S. Nagamine, C. Liu, M. Hashimoto, A. Minami and H. Oikawa, *Org. Lett.*, 2020, **22**, 1997–2001.

26 B. E. Cross, K. Norton and J. C. Stewart, *J. Chem. Soc. C*, 1968, 1054–1063.

27 K. Hanson, J. A. O'Neill, T. J. Simpson and C. L. Willis, *J. Chem. Soc., Perkin Trans. 1*, 1994, **17**, 2493–2497.

28 R. Schor, C. Schotte, D. Wibberg, J. Kalinowski and R. J. Cox, *Nat. Commun.*, 2018, **9**, 1963.

29 G. M. Strunz, M. Kakushima and M. A. Stillwell, *J. Chem. Soc., Perkin Trans. 1*, 1972, **18**, 2280–2283.

30 W. A. Ayer, P. P. Lu, H. Orszanska and L. Sigler, *J. Nat. Prod.*, 1993, **56**, 1835–1838.

31 V. Solovyev, P. Kosarev, I. Seledsov and D. Vorobyev, *Genome Biol.*, 2006, **7**, S10.

32 K. Rutherford, J. Parkhill, J. Crook, T. Horsnell, P. Rice, M.-A. Rajandream and B. Barrell, *Bioinformatics*, 2000, **16**, 944–945.

33 R. D. Finn, T. K. Attwood, P. C. Babbitt, A. Bateman, P. Bork, A. J. Bridge, H.-Y. Chang, Z. Dosztányi, S. El-Gebali, M. Fraser, J. Gough, D. Haft, G. L. Holliday, H. Huang, X. Huang, I. Letunic, R. Lopez, S. Lu, A. Marchler-Bauer, H. Mi, J. Mistry, D. A. Natale, M. Necci, G. Nuka, C. A. Orengo, Y. Park, S. Pesseat, D. Piovesan, S. C. Potter, N. D. Rawlings, N. Redaschi, L. Richardson, C. Rivoire, A. Sangrador-Vegas, C. Sigrist, I. Sillitoe, B. Smithers, S. Squizzato, G. Sutton, N. Thanki, P. D. Thomas, S. C. E. Tosatto, C. H. Wu, I. Xenarios, L.-S. Yeh, S.-Y. Young and A. L. Mitchell, *Nucleic Acids Res.*, 2017, **45**, D190–D199.

34 B. Tudzynski, M. Mihlan, M. C. Rojas, P. Linnemannstons, P. Gaskin and P. Hedden, *J. Biol. Chem.*, 2003, **278**, 28635–28643.

35 M. D. Lebar, J. W. Cary, R. Majumdar, C. H. Carter-Wientjes, B. M. Mack, Q. Wei, V. Uka, S. De Saeger and J. D. Di Mavungu, *Fungal Genet. Biol.*, 2018, **116**, 14–23.

36 P. Chankhamjon, D. Boettger-Schmidt, K. Scherlach, B. Urbansky, G. Lackner, D. Kalb, H.-M. Dahse, D. Hoffmeister and C. Hertweck, *Angew. Chem., Int. Ed.*, 2014, **126**, 13627–13631.

37 K. E. Lebe and R. J. Cox, *Chem. Sci.*, 2019, **10**, 1227–1231.

38 M. F. Grau, R. Entwistle, Y.-M. Chiang, M. Ahuja, C. E. Oakley, T. Akashi, C. C. C. Wang, R. B. Todd and B. R. Oakley, *ACS Chem. Biol.*, 2018, **13**, 3193–3205.

39 G. Yang, M. S. Rose, B. G. Turgeon and O. C. Yoder, *Plant Cell*, 1996, **8**, 2139–2150.

40 E. Geiser, S. K. Przybilla, A. Friedrich, W. Buckel, N. Wierckx, L. M. Blank and M. Böker, *Microb. Biotechnol.*, 2016, **9**, 116–126.

41 D. W. Brown, S.-H. Lee, L.-H. Kim, J.-G. Ryu, S. Lee, Y. Seo, Y. H. Kim, M. Busman, S.-H. Yun, R. H. Proctor and T. Lee, *Mol. Plant-Microbe Interact.*, 2015, **28**, 319–332.

42 L. Blank, J. Green and J. R. Guest, *Microbiology*, 2002, **148**, 133–146.

43 L. Eichinger, J. A. Pachebat, G. Glockner, M.-A. Rajandream, R. Suegang, M. Berriman, J. Song, R. Olsen, K. Szafranski, Q. Xu, B. Tunggal, S. Kummerfeld, M. Madera, B. A. Konfortov, F. Rivero, A. T. Bankier, R. Lehmann, N. Hamlin, R. Davies, P. Gaudet, P. Fey, K. Pilcher, G. Chen, D. Saunders, E. Sodergren, P. Davis, A. Kerhorne, X. Nie, N. Hall, C. Anjard, L. Hemphill, N. Bason, P. Farbrother, B. Desany, E. Just, T. Morio, R. Rost, C. Churcher, J. Cooper, S. Haydock, N. van Driessche, A. Cronin, I. Goodhead, D. Muzny, T. Mourier, A. Pain, M. Lu, D. Harper, R. Lindsay, H. Hauser, K. James, M. Quiles, M. M. Babu, T. Saito, C. Buchrieser, A. Wardrop, M. Felder, M. Thangavelu, D. Johnson, A. Knights, H. Loulseged, K. Mungall, K. Oliver, C. Price, M. A. Quail, H. Urushihara, J. Hernandez, E. Rabbinowitsch, D. Steffen, M. Sanders, J. Ma, Y. Kohara, S. Sharp, M. Simmonds, S. Spiegler, A. Tivey, S. Sugano, B. White, D. Walker, J. Woodward, T. Winckler, Y. Tanaka, G. Shaulsky, M. Schleicher, G. Weinstock, A. Rosenthal, E. C. Cox, R. L. Chisholm, R. Gibbs, W. F. Loomis, M. Platzer, R. R. Kay, J. Williams, P. H. Dear, A. A. Noegel, B. Barrell and A. Kuspa, *Nature*, 2005, **435**, 43–57.

44 B. Bonsch, V. Belt, C. Bartel, N. Duensing, M. Koziol, C. M. Lazarus, A. M. Bailey, T. J. Simpson and R. J. Cox, *Chem. Commun.*, 2016, **52**, 6777–6780.

45 A. Zajc, M. J. Romão, B. Turk and R. Huber, *J. Mol. Biol.*, 1996, **263**, 269–283.

46 J. I. Jimenez, J. F. Juarez, J. L. Garcia and E. Diaz, *Environ. Microbiol.*, 2011, **13**, 1718–1732.

47 F. Chardon and C. Damerval, *J. Mol. Evol.*, 2005, **61**, 579–590.



48 C. A. Cuomo, W. A. Untereiner, L.-J. Ma, M. Grabherr and B. W. Birren, *Genome Announc.*, 2015, **3**, e00021-00015.

49 Y.-C. Huang, Y.-H. Chen, S.-R. Lo, C.-I. Liu, C.-W. Wang and W.-T. Chang, *Mol. Microbiol.*, 2004, **53**, 81–91.

50 D. K. Holm, L. M. Petersen, A. Klitgaard, P. B. Knudsen, Z. D. Jarczynska, K. F. Nielsen, C. H. Gotfredsen, T. O. Larsen and U. H. Mortensen, *Chem. Biol.*, 2014, **21**, 519–529.

51 K. Katayama, T. Kobayashi, H. Oikawa, M. Honma and A. Ichihara, *Biochim. Biophys. Acta Protein Struct. Mol. Enzymol.*, 1998, **1384**, 387–395.

52 W. Kim, C.-M. Park, J.-J. Park, H. O. Akamatsu, T. L. Peever, M. Xian, D. R. Gang, G. Vandemark and W. Chen, *Mol. Plant-Microbe Interact.*, 2015, **28**, 482–496.

53 M. L. Nielsen, L. Albertsen, G. Lettier, J. B. Nielsen and U. H. Mortensen, *Fungal Genet. Biol.*, 2006, **43**, 54–64.

54 T. J. Carver, K. M. Rutherford, M. Berriman, M.-A. Rajandream, B. G. Barrell and J. Parkhill, *Bioinformatics*, 2005, **21**, 3422–3423.

55 M. S. Santisteban, T. Kalashnikova and M. M. Smith, *Cell*, 2000, **103**, 411–422.

56 G. Büchi, K. M. Snader, J. D. White, J. Zanos Gougoutas and S. Singh, *J. Am. Chem. Soc.*, 1970, **92**, 6638–6641.

57 T. Hosoe, J. B. Gloer, D. T. Wicklow, H. A. Raja and C. A. Shearer, *Heterocycles*, 2010, **81**, 2123–2130.

58 M. Isaka, M. Tanticharoen and Y. Thebtaranonth, *Tetrahedron Lett.*, 2000, **41**, 1657–1660.

