Published in final edited form as:

Nat Prod Rep. 2019 April 17; 36(4): 561–572. doi:10.1039/c8np00060c.

Nature and nurture: Confluence of pathway determinism with metabolic and chemical serendipity diversifies *Monascus* azaphilone pigments

Wanping Chen^{#1}, Yanli Feng^{#2}, István Molnár^{3,*}, and Fusheng Chen^{1,*}

¹Key Laboratory of Environment Correlative Dietology, College of Food Science and Technology, Huazhong Agricultural University, Wuhan, Hubei Province, 430070, China

²Hubei Key Laboratory of Edible Wild Plants Conservation and Utilization, College of Life Sciences, Hubei Normal University, Huangshi, Hubei Province, 435002, China

³Southwest Center for Natural Products Research, The University of Arizona, 250 E. Valencia Rd., Tucson, Arizona 85706, U.S.A.

[#] These authors contributed equally to this work.

Abstract

Understanding the biosynthetic mechanisms that generate the astounding structural complexity and variety of fungal secondary metabolites (FSMs) remains a challenge. As an example, the biogenesis of the *Monascus* azaphilone pigments (MonAzPs) has remained obscure until recently despite the significant medical potential of these metabolites and their long history of widespread use as food colorants. However, a considerable progress has been made in recent years towards the elucidation of MonAzPs biosynthesis in various fungi. In this highlight, we correlate a unified biosynthetic pathway with the diverse structures of the 111 MonAzPs congeners reported until June 2018. We also discuss the origins of structural diversity amongst MonAzPs analogues and summarize new research directions towards exploring novel MonAzPs. The case of MonAzPs illuminates the various ways that FSMs metabolic complexity emerges by the interplay of biosynthetic pathway determinism with metabolic and chemical serendipity.

Graphical Abstract

*Corresponding authors F. Chen, chenfs@mail.hzau.edu.cn, Tel.: 86-27-87282111, Fax: 86-27-87282111, I. Molnár, imolnar@email.arizona.edu, Tel: 1-520-621-9932, Fax: 1-520-621-8378.

⁶Conflicts of interest

There are no conflicts of interest to declare.

This highlight maps 111 congeners of *Monascus* azaphilone pigments on a unitary biosynthetic pathway and summarizes the causes for their diversity.

Keywords

Monascus; azaphilone pigments; food colorants; biosynthetic pathway; polyketides; metabolic complexity

1. Introduction

Monascus azaphilone pigments (MonAzPs) are a large group of secondary metabolites produced via the polyketide pathway by filamentous fungi, mainly by *Monascus* spp.. MonAzPs have been extensively used as natural food coloring agents for more than two thousand years, and they still play an important role in the worldwide food industry¹. For example, MonAzPs are traditional ingredients of sausages, fermented bean curd, red rice wine, *kamaboko*, and various other preserved dry meat and fish products in Asia. In Europe, MonAzPs had also been considered as partial substitutes for nitrate and nitrite salts for the preservation and color enhancement of meat products such as sausage, pâté, ham and frankfurters^{1, 2}. However, while MonAzPs are generally considered to be safe and even healthy in Asia, the widespread application of these food additives in Europe and the United States has been called into question after citrinin, a nephrotoxic, hepatotoxic and carcinogenic mycotoxin was detected in some batches of *Monascus* products². To provide MonAzPs products with acceptable safety profiles, many technological (fermentation and downstream processing) solutions have been devised and patented¹. These allow the suppression of citrinin production during industrial scale fermentation, and guarantee the manufacture of safe MonAzPs products, at least in China where maximum limit concentrations for citrinin in MonAzPs-related products are much lower than those in Europe and the United States^{3–5}. Thus, annual MonAzPs production is estimated to exceed 20,000 metric tons in China alone, and more than one billion people consume food containing MonAzPs-related products during their daily life⁶. Moreover, demand for MonAzPs as food additives is still growing rapidly due to their efficient and economical production on cheap substrates, their diverse colors and good solubility in water and ethanol⁷.

In addition to their widespread use as food colorants, MonAzPs also have many promising applications in the cosmetics, textile, printing and dyeing industries. For example, MonAzPs can be used as additives in cosmetics owing to their vivid color and excellent ability to absorb harmful near-ultraviolet light¹. They can also be utilized to produce ink for printers; to dye cotton yarn, leather and wool; and even to improve the efficiency of solar panels where MonAzPs are applied as a novel sensitizing dye in solar cells^{1, 8}.

MonAzPs products also possess a wide range of biological activities, including anticancer and antimutagenic, anti-diabetic, anti-inflammatory, antimicrobial, anti-hypercholesterolemic, and anti-obesity activities^{8, 9}. Herein, it is worth mentioning that *Monascus* spp. may produce many other bioactive metabolites along with MonAzPs. For instance, monacolin K, also known as lovastatin, is produced by *Monascus* spp. too. Lovastatin and its derivatives are widely used as serum cholesterol-lowering drugs for humans and animals due to their specific inhibition of 3-hydroxy-3-methylglutaryl-coenzyme A reductase, an enzyme that controls the rate of cholesterol biosynthesis¹⁰.

Despite their large-scale utilization and economic importance, our understanding of the biosynthesis of MonAzPs remains inadequate. Researches on MonAzPs started in the 1930s¹¹. In the 1970s, nuclear magnetic resonance analysis was used to successfully elucidate the structures of the six most abundant MonAzPs congeners. These “classical MonAzPs” (Supplementary Table 1) belong to the yellow (monascin **1** and ankaflavin **2**), orange (rubropunctatin **3** and monascorubrin **4**), and red pigment subclasses (rubropunctamine **5** and monascorubramine **6**), respectively⁸. By the end of 2011, about 50 MonAzPs were identified⁸. From 2012 to 2017 more than 50 additional MonAzPs compounds were discovered, thus bringing the total number of described MonAzPs to 111 (Supplementary Table 1). However, little progress has been made to decipher the biogenetic relationships of these compounds until recently. Thus, it remains puzzling whether a single biosynthetic pathway, or multiple pathways are behind the generation of the bewildering diversity and complexity of MonAzPs. Fortunately, a unifying picture of MonAzPs biosynthesis has emerged in the last 5 years as a result of genome sequencing, the discovery of MonAzPs biosynthetic gene clusters, and the functional analysis of some of the enzymes encoded in these clusters^{12, 13}. Such studies raised the possibility of further diversifying MonAzPs, and perhaps more importantly, also taught us about nature’s methods for diversity-oriented biosynthesis of fungal secondary metabolites (FSMs)¹⁴.

In the present highlight we summarize the 111 MonAzPs structures reported in the literatures until June 2018 (Supplementary Table 1) and map these congeners on a proposed unitary biosynthetic pathway. Clarification of the biogenetic relationships of the known MonAzPs congeners allows us to discern how pathway determinism and metabolic and chemical serendipity confer the observed MonAzPs diversity. Elucidation of these factors also allows us to discuss future directions to explore novel MonAzPs. We hope that this highlight not only contributes to a comprehensive understanding of MonAzPs biosynthesis, but also serves as a model towards exploring the generation of FSMs complexity in other fungi.

2. Functional analysis of MonAzPs biosynthetic gene clusters clarifies the trunk pathway

2.1 Organization of the MonAzPs gene clusters

Production of MonAzPs-type compounds is prevalent in the genera of *Monascus* and *Talaromyces*^{15, 16}, and several gene clusters responsible for the biosynthesis of these compounds have recently been identified from these fungi^{6, 12, 13, 17}. Comparison of these biosynthetic gene clusters shows that they share orthologous genes for a conserved MonAzPs trunk pathway (Fig. 1). However, the gene organization and (to a lesser extent) the gene content of this coevolving biosynthetic module are diversified in various fungal genomes by lineage-specific evolutionary events such as gene rearrangements, insertions of ancillary or even unrelated genes, and gene losses (Fig. 1)¹².

To date, five MonAzPs biosynthetic gene clusters are available from the *Monascus* genus: one each from *M. pilosus*¹³, *M. purpureus* YY-1⁶, *M. ruber* M7¹, *M. purpureus* NRRL1596 and *M. ruber* NRRL1597 (<http://genome.jgi.doe.gov/Monru1/Monru1.home.html>). These clusters all share the highly conserved regions I and II (Fig. 1) that encode the core enzymes for MonAzPs biosynthesis^{12, 13}. Genes in region III of *M. purpureus* NRRL1596, *M. purpureus* YY-1 and *M. ruber* M7, and those in region IV of *M. pilosus* and *M. ruber* NRRL1597 may not be necessary for MonAzPs assembly¹². In the genus *Talaromyces*, four MonAzPs clusters have been described to date: one each from *T. marneffei* ATCC18224, *T. stipitatus* ATCC10500, *T. aculeatus* ATCC10409 and *T. atrovirens* IBT11181, respectively^{18, 19}. The MonAzPs clusters in *T. marneffei* ATCC18224 and *T. stipitatus* ATCC10500 share an identical gene content and arrangement¹⁸. In *T. aculeatus* ATCC10409, the core cluster is supplemented by the insertion of a putative major facilitator superfamily (MFS) multidrug transporter-encoding gene. Curiously, the orthologous cluster in *T. atrovirens* IBT11181 is insufficient to encode MonAzPs biosynthesis on its own, as it lacks the crucial non-reducing polyketide synthase (nrPKS) responsible for the assembly of the first MonAzPs intermediate¹⁹. Thus, this cluster may have to rely on the supply of this intermediate from the mitorubrin nrPKS, encoded elsewhere in the *T. atrovirens* IBT11181 genome. If so, *T. atrovirens* IBT11181 represents an interesting example of secondary metabolic parsimony achieved by metabolic crosstalk²⁰.

2.2 Functional analysis of the MonAzPs gene clusters

The MonAzPs gene clusters of strains *M. ruber* M7 and *M. purpureus* KACC42430 have been the subjects of systematic functional analyses by gene knockouts, and in a few cases, heterologous expression and in vitro reconstitution of enzymatic reactions (Table 1).

Collectively, these studies reveal a unitary trunk pathway (Scheme 1) leading to the production of the classical pigments **1–4**, and the intermediates **11–12** and **38–39**. All other MonAzPs are generated by various shunt pathways branching off from the trunk pathway.

The trunk pathway of MonAzPs biosynthesis is initiated by a nrPKS (MrPigA in *M. ruber* M7) that features eight functional domains: a starter unit : acyl carrier protein transacylase (SAT), a ketoacyl synthase (KS), an acyltransferase (AT), a product template (PT), a C-

methyltransferase (MT), two acyl carrier proteins (ACP), and a reductive release domain (R). MrPigA produces the putative hexaketide benzaldehyde **P1**. The release of this very reactive intermediate from MrPigA may be aided by a deduced serine hydrolase (MrPigG in *M. ruber* M7). The first stable intermediate (**11**) is produced by a ketoreductase (MrPigC in *M. ruber* M7), which reduces the ω -1 carbonyl of compound **P1** to the alcohol. Next, **11** is hydroxylated at C-4 by a FAD-dependent monooxygenase (MrPigN in *M. ruber* M7) to yield the putative intermediate **P2**, triggering the formation of the pyran ring to afford compound **12**. The C-4 alcohol of the benzopyran in **12** is then acylated with a medium chain β -keto fatty acid by an acyltransferase (MrPigD in *M. ruber* M7) to yield the putative intermediate **P3**. The β -ketooctanoic or β -ketodecanoic acid moieties used for this acylation step are produced by a dedicated two-subunit fatty acid synthetase (MrPigJ and MrPigK in *M. ruber* M7) encoded in the MonAzPs gene clusters. The ω -1 alcohol of the putative acyl-benzopyran in **P3** then undergoes a two-step acetylation-elimination sequence. First, **P3** is acetylated at the O-11 position by an acetyltransferase (MrPigM in *M. ruber* M7) to afford the putative intermediate **P4**. Next, another acetyltransferase (MrPigO in *M. ruber* M7) eliminates acetic acid from **P4** to yield the putative intermediate **P5** with the C10(11) double bond. The C6(7) double bond of **P5** is reduced by a NAD(P)H-dependent oxidoreductase (MrPigE in *M. ruber* M7) to afford the putative intermediate **P6** that undergoes a spontaneous C5(2') Knoevenagel condensation to yield yellow pigments **38–39** with the characteristic linear tricyclic ring structure of the classical MonAzPs **1–4**. Starting from here, the trunk pathway bifurcates into two branches. The first branch leads to the classical yellow MonAzPs monascin **1** and ankaflavin **2** upon the reduction of the C5(2') double bond of **38** and **39** by a reductase (MrPigH and/or GME3457 that is encoded outside of the MonAzPs cluster in *M. ruber* M7). The other branch of the pathway yields the classical orange pigments rubropunctatin **3** and monascorubrin **4** by restoring the C6(7) double bond by a FAD-dependent oxidoreductase (MrPigF in *M. ruber* M7). Thus, the unitary MonAzPs trunk pathway only dictates the biosynthesis of the classical yellow and orange pigments **1–4** as its end products¹².

3. Shunt pathways branching from the trunk pathway generate diverse MonAzPs

In addition to the classical MonAzPs **1–6**, at least 105 natural MonAzPs congeners have been identified to date, including 47 red, 49 yellow, 8 orange and 1 purple pigments (Supplementary Table 1). Most of these compounds were isolated from *Monascus* spp. fermentations, although a few originate from *Talaromyces* (*Penicillium*) spp. cultures. However, the biogenetic relationships of these compounds have not been clarified in most cases. Now, the proposed MonAzPs trunk pathway (Scheme 1) allows us to map these compounds onto this unitary biosynthetic route, and to re-categorize them into 4 Stages based on their structural features and biogenetic relationships.

In the trunk pathway, the classical MonAzPs **1–4**, the intermediates **11**, **12**, **38** and **39**, and the putative intermediates **P1–P6** are produced (Scheme 1). The very large variety of other MonAzPs compounds may be generated from the highly reactive and unstable node

compounds of **P1-P6** which serve as the substrates for fortuitous metabolic pathways or uncatalyzed chemical reactions in the cell or in the culture broth.

3.1 MonAzPs 7–11 are Stage I compounds

Stage I compounds are hexaketides derived from the putative node compound **P1**. On the trunk pathway, reduction of **P1** by MrPigC or its orthologues yields compound **11**. Along **Shunt 1**, the release of **P1** from MrPigA or its orthologues is followed by a spontaneous aldol cyclization in the C-1 to C-10 register to yield the trihydroxynaphthalene MA-3 (**7**) (Scheme 2). Spontaneous oxidation of **7** yields benzoquinone MA-4 (**8**)^{12, 22} whose adventitious reduction affords MA-2 (**9**)¹². Alternatively, **7** can also be converted to MA-1 (**10**) by hydroxylation and subsequent reduction²².

3.2 MonAzPs 12 and 13 are Stage II compounds

Stage II compounds are pyrans without the characteristic fatty acyl side chains of MonAzPs, derived from putative node compound **P2**. On the trunk pathway, pyran ring closure affords compound **12** from **P2**. Monascusone A **13** is the only constituent along **Shunt 2**, derived by reduction of **12** in mutant *pigD-2* of *M. ruber* M7 (Scheme 3)¹².

3.3 MonAzPs 14–37 are Stage III compounds

MonAzPs **14-37** are Stage III compounds that are derived through **Shunts 3, 4 and 5** from three node compounds, the pyran fatty acyl esters **P3, P4** and **P5**, respectively. Stage III compounds typically undergo a C3–C2' Knoevenagel cyclization to yield an angular tricyclic carbon skeleton different from the linear tricyclic core of the classical MonAzPs **1–6**. Some of these compounds may be derived from partial biodegradation followed by additional modifications.

Shunt 3 branches out of the trunk pathway at **P3**(Scheme 4). Spontaneous intramolecular Knoevenagel cyclization of **P3** in the C3–C2' register yields MonAzPs **14–15**. Reduction of the C3(2') double bond of **14–15** affords monasfluols A-B (**16–17**)²⁹, while a subsequent reduction of the C-5 ketone produces monasfluol derivatives **18–19**²⁷. MonAzPs **20** likely forms by reduction of monasfluol A **16** followed by *O*-to-*N* substitution.

Shunt 4 affords the Stage III MonAzPs constituents **21–23** (Scheme 5). Spontaneous Knoevenagel cyclization of putative node compound **P4** in the C3–C2' register yields the putative intermediates **P7, P8** and **P9** that have not yet been characterized, probably due to their efficient conversion to acetyl-monasfluol A and B (**21** and **22**)¹² and monascuskaolin **23**³⁰ by the reduction of the C3(2') double bond. The C9 side chain featured in **23** is rare among MonAzPs.

Shunt 5 consists of various reactions that produce the Stage III MonAzPs **24–37** from the putative node compound **P5** (Scheme 6). Reduction and deacylation of **P5** yields FK17-P₂B₂ (**24**)¹². Cyclization of **24** with acetoacetic acid affords monascusone B **25**³¹. Alternatively, the chromophore of **24** may be desaturated at C6(7), followed by a direct reaction with lysine to replace the pyran oxygen with a nitrogen to afford compound **26**³² after further reduction. MonAzPs MC-2 (**27**), MC-4 (**28**), and the putative intermediates **P10** and **P11** are

formed when **P5** undergoes a spontaneous Knoevenagel cyclization in the C3–C2' register. These compounds are then further reduced to monasfluores A-B **29–30**¹², monascuskaodione **36** and monascuspurone **37**. Biodegradation of **29** by oxidative ring opening affords monapurone A **31**; the homologous compounds with alternative side chains, especially the expected C₇H₁₅ homologue, have not been identified yet. The enantiomeric pairs monapurones B/C (**32/33**) and monapurfluores A/B (**34/35**) are inferred to be formed by tetrahydrofuran ring closure and *O*-methylation of **31**.

3.4 MonAzPs 1–6 and 38–111 are Stage IV compounds

Stage IV MonAzPs are derived from putative node compound **P6** by C5–C2' Knoevenagel cyclization to yield the linear tricyclic carbon skeletons characteristic of the classical MonAzPs **1–6**. Stage IV compounds include the main products of the trunk pathway (**1–4**, **38–39**) and their derivatives along **Shunt 6** (**40–50**) and **Shunt 7** (**5–6**, **51–111**), including derivatives that have undergone partial biodegradation.

MonAzPs **40–50** are generated along **Shunt 6** that extends the trunk pathway starting with the yellow pigments **1–2** (Scheme 7). MonAzPs **47–48** derive from **1–2** by oxidative lactone ring opening. Sulfonation of these compounds yields **49–50**³³, while monaphilones B-A (**40–41**) result from decarboxylation of **47–48**. Compound **40** is further degraded to monaphilone C **42** by pyran ring opening and reduction of the double bond. A similar sequence of pyran ring opening and reduction may yield purpureusone **43** from **2**, without being prefaced by lactone ring opening. Finally, **1–2** may be converted to monascuspiloin **44** and monapilosusazaphilone **45** by reduction of the C-3' ketone to the corresponding alcohol. Monascusazaphilol **46** with the uncommon C9 side chain is likely generated in a similar manner from a still-unidentified homologue of **1** and **2**.

Starting from classical pigments **3** and **4**, **Shunt 7** extends the trunk pathway to afford MonAzPs **5–6** and **51–111** (Scheme 8). Many of these metabolites are amine derivatives that constitute the large family of red MonAzPs. Similar to **40–41** from **Shunt 6**, monarubrin **51** and rubropunctin **52** are produced by biodegradation of **3–4** by oxidative lactone ring opening, and then converted to **53–54** by *O*-to-*N* substitution and further redox reactions³⁴. Xanthomonasin A-B (**55–56**) may arise through oxidative ring opening of **3** and **4** followed by a complex series of successive rearrangements³⁵. Yellow II **57** may be afforded by oxidative chain shortening and reduction of the C1(2) double bond of **4**. Monapilols B and A (**58–59**) derive from **3** and **4** by reduction of the C-3 ketone, and can be further converted to monapilols D and C (**60–61**) and monasphilol-methoxy A and B (**62–63**) by forming acetone or methanol adducts, respectively³⁶. Monascopyridines A and B (**64–65**) may be produced from **5** and **6** by the reduction of the double bond at C5(2') followed by double-bond shift, and then further degraded to monascopyridines C and D (**66–67**) by oxidative lactone ring opening.

Compounds **5–6** and **68–93** are produced when MonAzPs **3–4** form adducts with various amines to effect an *O*-to-*N* substitution³⁷. Some of these adducts may further undergo reduction(s) and double bond migration at C1(2) and/or C6(7) to form MonAzPs

constituents **94-105**^{32, 38}. MonAzPs **106-107** are formed by reduction of the C5(2') double bond of **3-4**.

The origin of the C10(11) *cis* double bond in MonAzPs PP-O **108** and PP-Y **110** is unclear³⁹. The ω -carbon is oxidized to a carboxyl in **108**. Compounds PP-V **109** and PP-R **111** are derived from **108** and **110**, respectively, by *O*-to-*N* replacement after reaction with an amine⁴⁰.

4. The origins of MonAzPs diversity

Filamentous fungi can produce a huge number of secondary metabolites, many of which are of considerable interest from the medical, industrial, agricultural and economic standpoint⁴¹. The enormous structural diversity of FSMs is based on a relatively limited number of carbon skeletons produced by core enzymes such as polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs). However, these carbon skeletons are modified by a large number of tailoring enzymes whose expression is also modulated by the environmental conditions that fungi are exposed to^{42,42}. MonAzPs biosynthesis offers an ideal model system to investigate how structural variety and complexity result from FSMs biosynthetic pathways, and how such variety may be further expanded by combinatorial biosynthesis. Based on our current understanding, four key factors may be crucial for conferring the MonAzPs diversity.

4.1 Side chain length variability generates a series of MonAzPs homologues

The side chain β -keto fatty acyl moieties of MonAzPs are generated by a dedicated two-subunit fatty acid synthetase (FAS) (MrPigJ and MrPigK in *M. ruber* M7, **Fig. 2**) encoded within the pigment biosynthetic gene clusters¹². However, unlike their orthologues in aflatoxin/sterigmatocystin biosynthetic gene clusters⁴³, the chain length control of the MonAzPs FAS enzymes, and the substrate specificity of the acyltransferase (MrPigD in *M. ruber* M7) are somewhat relaxed. Thus, most MonAzPs feature either a β -ketooctanoic or a β -ketodecanoic acid moiety, but four MonAzPs constituents incorporate a β -keto fatty acyl moiety with a different chain length (C11 in **36**⁴⁴, C12 in **23**³⁰ and **46**⁴⁵, and C13 in **37**⁴⁶, Supplementary Table 1). It is likely that more such homologues will also be found in the future. Similarly, MonAzPs homologues with alternative side chains, including “unnatural” β -keto fatty acyls carrying various substituents and synthetic “handles” could conceivably be engineered by precursor feeding or other forms of combinatorial biosynthesis, tremendously increasing the MonAzPs structural diversity. Thus, imperfect pathway determinism by relaxed product specificity of the dedicated FAS and reduced substrate specificity of the cognate acyltransferase might increase MonAzPs diversity by affording a homologous series of pigments with various acyl side chains, and may allow the generation of novel MonAzPs.

4.2 Alternative outcomes during intramolecular Knoevenagel condensation diversify yellow MonAzPs

Spontaneous intramolecular Knoevenagel aldol condensation plays a key role in MonAzPs biosynthesis, as it initiates **Shunts 3-5** towards various yellow MonAzPs (Scheme 1). This condensation has two facile regioisomeric outcomes: formation of an angular tricyclic

carbon skeleton in the C3–C2' register, thermodynamically favored in the early intermediates **P3** to **P5**; and creation of a linear tricyclic ring system in the C5–C2' register, preferred in intermediate **P6** (Scheme 9). Despite the crucial importance of this reaction to generate the characteristic ring system of MonAzPs, the regiospecificity of this aldol condensation is only indirectly controlled by the enzymes encoded in the MonAzPs pathways. This is presumably achieved by kinetic control of the reactions along the trunk pathway from putative intermediate **P3** to **P6**, and eventually by modulating the electrophilicity of the C-5 ketone of the acyl benzopyran **P6** by breaking the π -conjugated system in **P5**^{12, 22}. However, this control over the regiospecificity of the Knoevenagel condensation is by no means absolute, and representative angular tricyclic shunt products such as **16**, **17**, **29** and **30** also accumulate in wild type strains in addition to the main linear tricyclic MonAzPs¹¹. Conversely, strains in which the trunk pathway has been blocked by knocking out the *mrpigM*, *mrpigO* or *mrpigE* genes accumulate angular tricyclic derailment products along **Shunt 3**, **4** and **5** (Table 1) almost exclusively, with only minor amounts of products detectable that originated from Knoevenagel condensation in the C5–C2' register. This is even more remarkable considering that a C5–C2' condensation in **P5** would directly generate the classical orange pigments **3** and **4**, thus shortening the biosynthetic pathway (Scheme 9). An analogous bias for regiospecific cyclization has also been observed during the biosynthesis of chaetoviridins, azaphilones produced by *Chaetomium* spp.²⁹. Thus, incomplete pathway determinism by imperfect enzymatic control over Knoevenagel condensation of the reactive intermediates along the trunk pathway is an important factor for the serendipitous generation of structural diversity amongst the yellow MonAzPs.

4.3 Uncatalyzed reactions with various amine substrates diversify red MonAzPs

Generally speaking, MonAzPs are stable under various conditions, including changes in temperature, pH, light, metal ions and so on⁸, but the pyran ring of the classical orange pigments **3** and **4** is prone to spontaneous *O*-to-*N* substitution by direct reaction with endogenous amines to afford the γ -vinylogous pyridines of the red MonAzPs^{47, 48}. As shown in Supplementary Table 1, numerous red pigments originate by reaction with amino acids (including both D- and L-amino acids), and *N*-glucosyl derivatives have also been isolated⁴⁹. It is believed that there are many more red MonAzPs still waiting to be isolated and characterized. *In vitro* chemical reactions indicated that ammonia or basic amino acids such as arginine are the preferred nucleophiles, and that alkaline conditions under which the amine group is not protonated would promote *O*-to-*N* substitution¹². Collectively, the facile non-enzymatic reaction between the orange MonAzPs **3–4** that are two of the end products of the trunk pathway, and the wide variety of amines available in the cells or in the fermentation media is the main cause of the wide diversity of red MonAzPs. Consequently, judicious choice of nitrogen sources for the production media can yield more “pure” (chemically uniform) red MonAzPs⁵⁰. And even during the MonAzPs extraction process, we think that if some special ammonium compounds are added, the relative novel MonAzPs compounds may be appeared, although they should be precisely considered as artefacts, not natural products, and up to now, no reference has been published about this.

4.4 Serendipitous enzymatic reactions increase MonAzPs diversity

All the shunt pathways described in this highlight also feature serendipitous enzymatic reactions, especially oxidations and reductions that are catalyzed by host enzymes not encoded in the MonAzPs biosynthetic gene clusters. The identification of the responsible enzymes and their genes represents a new frontier for MonAzPs research, but is far from trivial due to the large number of potentially relevant enzymes encoded in the fungal genomes. In addition, these biotransformations vary by MonAzPs producers, and even analogous reactions may be carried out by unrelated enzymes in different strains.

While most of these biotransformation reactions currently remain cryptic, a candidate enzyme has been identified for the reduction of the C3(2') double bond in various MonAzPs shunt products. This enzyme, GME3457 of *M. ruber* M7, is an orthologue of CazP, a dehydrogenase encoded within the chaetoviridin biosynthetic cluster, which shows considerable substrate promiscuity and contributes to the chemical diversity of chaetoviridins in *Chaetomium* spp. (Scheme 10A)²⁹. Key MonAzPs intermediates such as **14**, **15**, **27**, **28**, and **P7-P11** display similar carbon skeletons to chaetoviridin H, a substrate of CazP (Scheme 10B). While these compounds are necessary intermediates for many angular tricyclic shunt products, they have only been detected in trace amounts or were missing from extracts of *Monascus* fermentations¹². Just as with the analogous intermediate chaetoviridin H^{11, 28}, the low abundances of these intermediates could be explained by the efficient reduction of their C3(2') double bonds, presumably by a CazP-like enzyme (Scheme 10). Searching the encoded proteome of *M. ruber* M7 with CazP as the bait identified the plausible orthologue GME3457 (GenBank accession: KY270501)¹², and orthologous enzymes are also encoded by the other *Monascus* and *Talaromyces* MonAzPs producers. Gene knockout and/or heterologous expression and reconstitution of the enzymatic activity will reveal whether GME3457 and its orthologues are indeed responsible for the reduction of the C3(2') double bond in many MonAzPs shunt products. Additional enzymes in the various MonAzPs producers accomplish many other serendipitous biotransformations including partial breakdown of the tricyclic pigment structure, and thus enrich the diversity of MonAzPs congeners.

5. Conclusion and prospects

In this highlight, we summarized the results of recent studies on the biosynthesis of MonAzPs in several Eurotiomycete fungi, collected the currently known MonAzPs congeners whose structures are described in the literatures (Supplementary Table 1), and attempted to map these structures on a consensus biosynthetic pathway. Although the structures of MonAzPs are extremely diverse and often highly complex, their biosynthesis employs a unitary trunk pathway that ushers intermediates towards the classical yellow (**1**, **2**) and orange (**3**, **4**) pigments, and features a variety of shunt pathways branching off from the trunk pathway at highly reactive node compounds. We classified the 111 known MonAzPs into four Stages based on their structural features and deduced biogenesis. We have also identified four strategies that nature employs to generate the astonishing structural variety of these pigments by using simple and achiral starting blocks. First, the MonAzPs biosynthetic machinery employs key enzymes with moderately relaxed product specificities, as

exemplified by the FAS and the acyltransferase that generate and attach a small cohort of β -keto fatty acyl side chains to the polyketide core in the various pigment congeners. Second, the MonAzPs pathway exerts an effective but still limited control over the regiospecificity of the spontaneous intramolecular Knoevenagel cyclization that generates the tricyclic core of most MonAzPs. Failure in this control shunts intermediates towards MonAzPs with an angular, as opposed to a linear, tricyclic ring system. Third, uncatalyzed *O*-to-*N* substitution reactions with available amines in the cell or the culture media generate a large variety of red MonAzPs with a γ -vinylogous pyridine core. Finally, serendipitous reactions with enzymes encoded outside of the MonAzPs biosynthetic gene cluster on the fungal genome further diversify MonAzPs congeners in all shunt pathways. Similar principles are no doubt also involved in many other secondary metabolic pathways, especially in fungi that often elaborate a relatively few simple carbon skeletons from polyketide biosynthesis to produce a large variety of complex secondary metabolites.

We expect that a better understanding of the MonAzPs pathway will continue to emerge by comparative genomics, classical gene knockouts, and increasingly, by heterologous expression and enzyme or pathway reconstitution. Improved detection, isolation and structure elucidation methods will be similarly useful to continue to map the natural and engineered diversity of MonAzPs. In addition, combinatorial biosynthetic methods are expected to generate unprecedented, unnatural MonAzPs analogues and hybrid, chimeric metabolites using enzymes from divergent biosynthetic pathways. Meanwhile, biosynthesis will be directed towards the MonAzPs congeners that are most desirable for the various food, industrial and health care applications, and metabolic engineering will enhance the safety and the productivity of the fermentations. Thus, modern technologies will continue to enhance the economic impact of MonAzPs that have been known and used since antiquity. At the same time, lessons learned about the generation of metabolic complexity by “nature and nurture”, that is, encoded pathway determinism and metabolic/chemical serendipity, will also help us understand and exploit fungal secondary metabolism in general.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors express special thanks to laboratory members, especially Nana Xie, Qingpei Liu, Jiao Liu, Mei Zhang, Runfa Chen and Wang Chen, for their continuous efforts to clarify the MonAzPs biosynthetic pathway. This work was supported by the Major Program of National Natural Science Foundation of China (Nos. 31330059 and 31730068 to F.C.), the National Key Research and Development Program of China (No. 2018YFD0400404 to F.C.), the National Natural Science Foundation of China (No. 31601446 to W.C.), the Natural Science Foundation of Hubei Province in China (No. 2017CFB367 to Y.F.), the U.S. National Institutes of Health (NIGMS R01GM114418 to I. M.) and the Higher Education Institutional Excellence Program of the Ministry of Human Capacities in Hungary (Biotechnology thematic program to I. M.). I. M. has disclosed financial interests in TEVA Pharmaceutical Industries Ltd. (Hungary) and DSM Nutritional Products LLC (USA), which are unrelated to the subject of the research presented here.

8. References

1. Chen W, He Y, Zhou Y, Shao Y, Feng Y, Li M and Chen F, Compr. Rev. Food Sci, 2015, 14, 555–567.

2. Mapari SAS, Thrane U and Meyer AS, Trends Biotechnol, 2010, 28, 300–307. [PubMed: 20452692]
3. National Food Safety Standards GB1886.181–2016, 2016, 1–5.
4. Commission Regulation (EU) No 212/2014, 2014, 3–4.
5. Gupta RC, Lasher MA, Mukherjee IRM, Srivastava A and Lall R, in Reproductive and Developmental Toxicology (Second Edition), ed. Gupta RC, Academic Press, 2017, DOI: 10.1016/B978-0-12-804239-7.00048-2, pp. 945–962.
6. Yang Y, Liu B, Du XJ, Li P, Liang B, Cheng XZ, Du LC, Huang D, Wang L and Wang S, Sci. Rep, 2015, 5, 8331. [PubMed: 25660389]
7. Mostafa ME and Abbady MS, Global J. Biotechnol. Biochem, 2014, 9, 1–13.
8. Feng YL, Shao YC and Chen FS, Appl. Microbiol. Biot, 2012, 96, 1421–1440.
9. Lee CL, Wen JY, Hsu YW and Pan TM, J. Microbiol. Immunol, 2018, 51, 27–37.
10. Lachenmeier DW, Monakhova YB, Kuballa T, Lobell-Behrends S, Maixner S, Kohl-Himmelseher M, Waldner A and Steffen C, Chin. Med, 2012, 7. [PubMed: 22417176]
11. Salomon H and Karrer P, Helv. Chim. Acta 1932, 15, 18–22.
12. Chen W, Chen R, Liu Q, He Y, He K, Ding X, Kang L, Guo X, Xie N and Zhou Y, Chem. Sci, 2017, 8, 4917–4925. [PubMed: 28959415]
13. Balakrishnan B, Karki S, Chiu SH, Kim HJ, Suh JW, Nam B, Yoon YM, Chen CC and Kwon HJ, Appl. Microbiol. Biot, 2013, 97, 6337–6345.
14. Helaly SE, Thongbai B and Stadler M, Nat. Prod. Rep, 2018, DOI: 10.1039/C8NP00010G.
15. Frisvad JC, Yilmaz N, Thrane U, Rasmussen KB, Houbraiken J and Samson RA, Plos One, 2013, 8, e84102. [PubMed: 24367630]
16. Shao Y, Lei M, Mao Z, Zhou Y and Chen F, Appl. Microbiol. Biot, 2014, 98, 3911–3922.
17. Xie N, Liu Q and Chen F, Biotechnol. Lett, 2013, 35, 1425–1432. [PubMed: 23690031]
18. Nierman WC, Fedorova-Abrams ND and Andrianopoulos A, Genome A, 2015, 3, e01559–01514. [PubMed: 25676766]
19. Rasmussen KB, Ph.D, Technical University of Denmark, 2015.
20. Wiemann P, Guo CJ, Palmer JM, Sekonyela R, Wang CCC and Keller NP, P. Natl. Acad. Sci. USA, 2013, 110, 17065–17070.
21. Xie N, Zhang Y and Chen F, Acta Microbiologica Sinica, 2015, 55, 863–872. [PubMed: 26710605]
22. Balakrishnan B, Suh JW, Park SH and Kwon HJ, RSC Adv, 2014, 4, 59405–59408.
23. Liu Q, Xie N, He Y, Wang L, Shao Y, Zhao H and Chen F, Appl. Microbiol. Biot, 2014, 8, 285–296.
24. Balakrishnan B, Park SH and Kwon H-J, Appl. Biol. Chem, 2017, 60, 437–446.
25. Balakrishnan B, Chandran R, Park SH and Kwon HJ, J. Microbiol. Biotechnol, 2015, 25, 1648–1652.
26. Balakrishnan B, Park SH and Kwon HJ, Biotechnol. Lett, 2017, 39, 163–169. [PubMed: 27714556]
27. Balakrishnan B, Kim HJ, Suh JW, Chen CC, Liu KH, Park SH and Kwon HJ, J. Korean Soc. Appl. Bi, 2014, 57, 191–196.
28. Balakrishnan B, Chen CC, Pan TM and Kwon HJ, Tetrahedron Lett, 2014, 55, 1640–1643.
29. Sato M, Winter JM, Kishimoto S, Noguchi H, Tang Y and Watanabe K, Org. Lett, 2016, 18, 1446–1449. [PubMed: 26959241]
30. Cheng MJ, Wu MD, Su YS, Yuan GF, Chen YL and Chen IS, Phytochem. Lett, 2012, 5, 262–266.
31. Jongrungruangchok S, Kittakoop P, Yongsmith B, Bavovada R, Tanasupawat S, Lartpornmatulee N and Thebtaranonth Y, Phytochemistry, 2004, 65, 2569–2575. [PubMed: 15451319]
32. Lian XJ, Wang CL and Guo KL, Dyes Pigments, 2007, 73, 121–125.
33. Yang H, Li J, Wang Y and Gan C, Food Chem, 2018, 245, 536–541. [PubMed: 29287406]
34. Loret MO and Morel S, J. Agric. Food Chem, 2010, 58, 1800–1803. [PubMed: 20030327]
35. Sato K, Iwakami S, Goda Y and Okuyama E, Heterocycles, 1992, 34, 2057–2060.
36. Huang Z, Zhang S, Xu Y, Li L and Li Y, Phytochem. Lett, 2014, 10, 140–144.

37. Hajjaj H, Klaébé A, Goma G, Blanc PJ, Barbier E and Francoeois J, *Appl. Environ. Microbiol*, 2000, 66, 1120–1125. [PubMed: 10698780]
38. Izawa S, Harada N, Watanabe T, Kotokawa N, Yamamoto A, Hayatsu H and Arimoto-Kobayashi S, *J. Agric. Food Chem*, 1997, 45, 3980–3984.
39. Ogihara J and Oishi K, *J. Biosci. Bioeng*, 2002, 93, 54–59. [PubMed: 16233165]
40. Mapari SAS, Hansen ME, Meyer AS and Thrane U, *J. Agric. Food Chem*, 2008, 56, 9981–9989. [PubMed: 18841978]
41. Wiemann P and Keller NP, *J. Ind. Microbiol. Biot*, 2014, 41, 301–313.
42. Schor R and Cox R, *Nat. prod. rep*, 2018, 35, 230–256. [PubMed: 29537034]
43. Yu J, *Toxins*, 2012, 4, 1024–1057. [PubMed: 23202305]
44. Cheng MJ, Wu MD, Chan HY, Cheng YC, Chen JJ, Chen YL, Chen IS and Yuan GF, *Chem. Nat. Compd*, 2015, 51, 1091–1093.
45. Cheng MJ, Wu MD, Yuan GF, Su YS and Yanai H, *Phytochem. Lett*, 2012, 5, 567–571.
46. Cheng MJ, Wu MD, Chan HY, Chen JJ, Cheng YC, Chen YL, Chen IS and Yuan GF, *Chem. Nat. Compd*, 2016, 52, 231–233.
47. Lin TF, Yakushijin K, Bt GH~chi and A. L. Demain, *J. Ind. Microbiol*, 1992, 9, 173–179.
48. Martinkova L, *Food Addit. Contam*, 1999, 16, 15–24. [PubMed: 11565571]
49. Hajjaj H, Klaébé A, Loret MO, Tzedakis T, Goma G and Blanc PJ, *Appl. Environ. Microbiol*, 1997, 63, 2671–2678. [PubMed: 16535644]
50. Shi K, Song D, Chen G, Pistolozzi M, Wu Z and Quan L, *J. Biosci. Bioeng*, 2015, 120, 145–154. [PubMed: 25648278]

Fig.1. MonAzPs biosynthetic gene clusters from the genera of *Monascus* and *Talaromyces*.

A. Comparison of the structures of MonAzPs gene clusters among *Monascus* and *Talaromyces* strains. Orthologous genes are highlighted by the same colors and assigned the same one-letter code. **B.** Biosynthetic and ancillary proteins encoded in the *Monascus* and *Talaromyces* MonAzPs clusters. Annotations of the encoded proteins A-P are based on the *M. ruber* M7 nomenclature. Numbers show the percent identities of the orthologous proteins in other strains as compared with those in *M. ruber* M7. Proteins encoded by genes absent in *M. ruber* M7: R, cytochrome P450; S, MFS transporter; A1, hypothetical protein; A2, GTP-binding protein; A3, transcription factor; A4, Kelch repeat protein; A5, iron-regulated transporter; A6, hypothetical protein. —, Genes encoding similar proteins are not present in the gene cluster. This figure was adapted from Chen et al., Chem. Sci., 2017, 8, 4917–4925 – published by The Royal Society of Chemistry - with some modification¹².

Scheme 2.
Shunt 1 yields the Stage I compounds 7–10

Scheme 3.
Shunt 2 yields the Stage II compound 13

Scheme 4.
Shunt 3 yields the Stage III compounds 14–20

Scheme 5.
Shunt 4 yields the Stage III MonAzPs 21–23

Scheme 6.
Shunt 5 yields the Stage III MonAzPs 24–37

Scheme 7.
Shunt 6 yields the Stage IV MonAzPs 40–50

Scheme 8. Shunt 7 yields the Stage IV MonAzPs 5–6, 51–111.

Additional structures are shown in Supplementary Table 1.

Scheme 9. Regioselectivity of Knoevenagel condensations during MonAzPs biosynthesis.
Red crosses indicate that these reactions are seldom if ever observed.

Scheme 10. A CazP-like enzyme for the reduction of key azaphilone intermediates.

A. Substrate promiscuity of dehydrogenase CazP during chaetoviridin production in *Chaetomium* spp.²⁹ **B.** Reduction of MonAzPs intermediates may be catalyzed by the CazP-like enzyme GME3457 in *M. ruber* M7.

Table 1Functions of MonAzPs gene clusters in *M. ruber* M7 and *M. purpureus* KACC42430

<i>M. ruber</i> M7 ¹	<i>M. purpureus</i> KACC42430 ²	Proof ³	Results	References
Biosynthetic genes				
<i>mrpA</i>	<i>MpPKS5</i>	KO HE	Abolishment of MonAzPs production Accumulation of Shunt 1 products	12, 13, 21
<i>mrpC</i>	<i>mppA</i>	KO HE	Accumulation of Shunt 1 compounds Reconstituted with <i>mrpA</i> + <i>mrpG</i> : Accumulation of compound 11	12, 22
<i>mrpD</i>	<i>mppB</i>	KO IVR	Accumulation of compounds 12,13 Production of P3 analogues	12
<i>mrpE</i>	<i>mppC</i>	KO	Accumulation of Shunt 5 compounds	22, 23
<i>mrpF</i>	<i>mppG</i>	KO	No orange or red pigments, accumulation of 1 and 2	12, 24
<i>mrpG</i>	<i>mppD</i>	KO HE	Reduced MonAzPs yield Reconstituted with <i>mrpA</i> : Accumulation of Shunt 1 products	12, 25
<i>mrpH</i>	<i>mppE</i>	KO	Enhanced orange and red pigment production	26
<i>mrpJ</i>	<i>MpFasA2</i>	KO	Accumulation of compounds 12,13	12
<i>mrpK</i>	<i>MpFasB2</i>	KO	Accumulation of compounds 12,13	12, 27
<i>mrpM</i>	<i>mpp7</i>	KO	Accumulation of Shunt 3 compounds	12, 28
<i>mrpN</i>	<i>mppF</i>	KO	Accumulation of compound 11	12, 22
<i>mrpO</i>	<i>mpp8</i>	KO	Accumulation of Shunt 4 compounds	12
Transport and regulation genes				
<i>mrpB</i>	<i>mppR1</i>	KO	Abolishment of MonAzPs production	13, 17
<i>mrpI</i>	<i>mppR2</i>	-	No functional verification so far	
<i>mrpL</i>	<i>mpp15</i>	KO	Reduced MonAzPs yield	12
<i>mrpP</i>	<i>mppI</i>	KO	Accumulation of intracellular yellow pigments	12

¹ *M. ruber* M7, gene names in the MonAzPs biosynthetic cluster of *M. ruber* M7.² *M. purpureus* KACC42430, gene names in the MonAzPs biosynthetic gene cluster of *M. purpureus* KACC42430.³ Proof, the nature of the functional studies conducted for the genes. KO, gene knockout; IVR, in vitro reconstitution of the enzyme; HE, heterologous expression.