Divergence of metabolites in three phylogenetically close *Monascus* species (*M. pilosus*, *M. ruber*, and *M. purpureus*) based on secondary metabolite biosynthetic gene clusters

Yuki Higa¹, Young-Soo Kim¹, Md. Altaf-Ul-Amin², Ming Huang², Naoaki Ono³,²* and Shigehiko Kanaya²,³

**Abstract**

**Background:** Species of the genus *Monascus* are considered to be economically important and have been widely used in the production of yellow and red food colorants. In particular, three *Monascus* species, namely, *M. pilosus*, *M. purpureus*, and *M. ruber*, are used for food fermentation in the cuisine of East Asian countries such as China, Japan, and Korea. These species have also been utilized in the production of various kinds of natural pigments. However, there is a paucity of information on the genomes and secondary metabolites of these strains. Here, we report the genomic analysis and secondary metabolites produced by *M. pilosus* NBRC4520, *M. purpureus* NBRC4478, and *M. ruber* NBRC4483, which are NBRC standard strains. We believe that this report will lead to a better understanding of red yeast rice food.

**Results:** We examined the diversity of secondary metabolite production in three *Monascus* species (*M. pilosus*, *M. purpureus*, and *M. ruber*) at both the metabolome level by LCMS analysis and at the genome level. Specifically, *M. pilosus* NBRC4520, *M. purpureus* NBRC4478 and *M. ruber* NBRC4483 strains were used in this study. Illumina MiSeq 300 bp paired-end sequencing generated 17 million high-quality short reads in each species, corresponding to 200 times the genome size. We measured the pigments and their related metabolites using LCMS analysis. The colors in the liquid media corresponding to the pigments and their related metabolites produced by the three species were very different from each other. The gene clusters for secondary metabolite biosynthesis of the three *Monascus* species also diverged, confirming that *M. pilosus* and *M. purpureus* are chemotaxonomically different. *M. ruber* has similar biosynthetic and secondary metabolite gene clusters to *M. pilosus*. The comparison of secondary metabolites produced also revealed divergence in the three species.

(Continued on next page)
Background

Species of the genus *Monascus* are economically important because they have been widely used in the production of yellow and red food colorants. In particular, *M. pilosus*, *M. purpureus*, and *M. ruber* are commonly used for food fermentation in the cuisine of East Asian countries including China, Japan, and Korea [1–3]. These species are also utilized to produce various kinds of natural pigments (reviewed in [4, 5]), including yellow pigments such as ankaflavin, monascin and rubropunctatin, orange pigments such as monascorubrin, purple pigments such as rubropunctamin and monascorubramin, and red pigments such as monascorubramine, *N*-glucose-7-monascorubramine, *N*-glucosylmonascorubramine, *N*-glutaryl-monascorubramine, and *N*-glutaryl-7-monascorubramin. One example is the traditionally fermented rice that contains at least 6 pigments from *Monascus* spp., including rubropunctatin, monascorubrin, rubropunctamin, monascorubramin, ankaflavin, and monascin [6]. The prediction of biosynthetic pathways for structurally diverse azaphilone pigments has recently been reported [7]. Recently, azaphilone pigment has been found to be produced by *Penicillium marneffei* and *Talaromyces atroroseus* [8, 9].

*M. pilosus* is a well-known fungus that produces several bioactive metabolites, such as monacolins K and L, as well as several pigments that are related with biological activities including anti-obesity, regulation of lipid metabolism, and Alzheimer’s disease at the in vitro and in vivo levels [8]. *M. purpureus* contains unsaturated fatty acids, sterols, monacolin and azaphilone pigments. It has been reported that these compounds are effective in lowering cholesterol, as well as in the treatment of diabetes, cardiovascular diseases, and some cancers [10]. *M. ruber* contains monacolin and azaphilone pigment. Recent studies have investigated *M. ruber* as an alternative to nitrite substitutes in meat processing [11].

The complete genome sequence of the industrial strain *M. purpureus* YY-1 is already available [12]. Here, we determined the draft genome sequences of *M. ruber* and *M. pilosus* in order to compare their different phenotypes. Understanding the diversity of the secondary metabolites produced by these species at the genome level is crucial for their industrial applications. We analyzed *M. pilosus* NBRC4520, *M. purpureus* NBRC4478, and *M. ruber* NBRC4483 to determine the diversity of the pigments based on metabolome data and secondary metabolite-related gene clusters such as monacolins, citrinin and azaphilone pigment. Several pigments are synthesized by PKS (poly ketide synthase) enzyme and NRPS (nonribosomal peptide) enzyme systems, which are encoded by large gene clusters in the genome. Comparison of such gene clusters between the three species will provide new insights into the potential production of novel pigments.

*Monascus* species produce a multitude of compounds, including polyketides, unsaturated fatty acids, and phytosterols. Monacolins, especially monacolin K, inhibit 3-hydroxy-3-methylglutaryl-coenzyme A reductase, which is the rate-limiting step in cholesterol biosynthesis. These compounds found in red yeast rice prevent high cholesterol levels that causes atherosclerosis [13]. Hence, it is expected that metabolites related to *Monascus* pigments can contribute to human health. However, citrinin was found as an undesirable contaminant in red yeast rice [14]. Specifically, citrinin has been reported to be nephrotoxic and must be strictly controlled.

Thus, it is important to define the diversity of biosynthetic pathways responsible for secondary metabolite production in economically important species such as *Monascus*. Red yeast rice is used in many foods around the world. Foods using red yeast rice must contain substances such as monacolin K that contribute to health, but must not contain citrinin, which causes nephrotoxicity. Therefore, further studies of genes involved in the synthesis of secondary metabolites by *Monascus* need to be carried out. In the present study, we determined the genome sequences of *M. pilosus*, *M. purpureus*, and *M. ruber*. The phylogenetic and chemotaxonomic differences between these three species were characterized by analyzing the gene clusters associated with secondary metabolites.

Results

Cultivation of *Monascus* and LCMS analysis

*Monascus* species can produce several types of azaphilones, including nitrogenated azaphilones, *N*-glucosyl azaphilones, amino acid derivative azaphilones, and citrinins [4]. We cultivated the three *Monascus* species to compare the production of the pigments and their
related metabolites using both agar and liquid medium, i.e., potato dextrose agar (PDA) and potato dextrose liquid (PDL) medium, which is the most frequently used culture medium for Monascus growth and metabolite production [15]. As shown in Fig. 1a, distinct colony shapes and colors could be observed on PDA and in PDL among M. pilosus, M. ruber, and M. purpureus. In order to quantify the difference of the pigment contents, we analyzed the medium using LC-MS and identified 14 pigments in total. Figure 1b showed the concentration of 14 metabolites quantified in three biological replicates in the three species as a heatmap with two-dimensional hierarchical clustering to display their similarity. This result confirmed that not all these metabolites are synthesized in the three species. The numbers of pigments commonly identified in each species are summarized as the Venn diagram (Fig. 1c). Dehydromonacolin K, rubropunctatin, monascin, and ankaflavin 2 were commonly produced by all three Monascus species. Of the three species, M. pilosus produced the greatest number of pigments (12 in total; Fig. 1c). Ten pigments, except monascorubramine, were produced by M. pilosus, while ankaflavin 1 and rubropunctamine were only produced by M. pilosus. Citrinin, a mycotoxin with nephrotoxic activity in mammals [16], was only produced by M. purpureus.

**Biosynthetic pathway of detected metabolites**

In order to understand the difference of the metabolites produced in these three species, we investigated their biosynthesis pathways in detail and illustrated in Fig. 2. It has been identified that the biosynthesis of the precursor of these metabolites, 1H-isochromenes are started from malonyl-CoA, in Penicillium marneffei and M. ruber [17, 18]. Citrinin polyketide synthase (PKS) converts the PKS-bound product citrinin [17]. The biosynthetic pathway from malonyl-CoA to monacolins was also determined in Aspergillus terreus, M. ruber, and M. purpureus [19, 20]. It should be noted that PKS-bound products are acted upon by two different types of PKS enzymes – one is an enzyme to produce Monascus azaphilone pigments, which corresponds to the pathway from malonyl-CoA to 1H-isochromenes and nitrogenated azaphilones [17], and the other is citrinin polyketide synthase, which corresponds to the pathway from

---

**Fig. 1** Pigments for three Monascus species for PD agar and PDL cultures. 

- **a** Monascus species are cultured in PD agar and PDL medium at 30 °C.
- **b** 2D clustering of pigment quantities among the three species.
- **c** Venn diagram of the pigments observed in PDL culture. Two-letter abbreviations used for the Venn diagrams: Pi, M. pilosus NBRC 4520; Ru, M. ruber NBRC 4483; Pu, M. purpureus NBRC 4478.
PKS-bound product to citrinin [21]. We also added the amount of the metabolites associated with each precursor as color bars. Among the three *Monascus* species, all eight metabolites related to 1H-isochromenes were only detected in *M. pilosus*. As for the other pathways, while citrinin was observed only in *M. purpureus*, Dehydromonacolin K, which is a precursor for monacolin K production, was detected in all three species. These results showed that the production of secondary metabolites are distinct among these three *Monascus* species.

**Diversity and classification of secondary metabolites produced by the three *Monascus* species**

Fungal polyketides are the largest and most structurally diverse type of secondary metabolites, ranging from simple aromatic compounds to complex macrocyclic lactones [4]. Azaphilones are pigments with pyrone-quinone structures containing a highly oxygenated bicyclic core and a chiral quaternary center and produced mainly by *Monascus* species [21]. To examine the metabolite divergence in the three *Monascus* species, we examined the diversity of azaphilones in the abovementioned five major metabolic groups. Metabolites belonging to these groups are synthesized by PKS enzymes [19, 20]. Similar studies have been performed for 373 fungi azaphilones [4], monacolins [5, 22–25], and monaphilones [26]. We examined a total of 74 secondary metabolites (Fig. 3a) and identified the ubiquitous and species-specific metabolites in the five groups using Venn diagrams. Results showed that the three *Monascus* species can produce unique 1H-isochromene compounds. *M.
ruber and M. purpureus can produce a large variety of secondary metabolites, while individually, M. pilosus, M. ruber, and M. purpureus can produce a large variety of nitrogenated azaphilones, monacolins, and 1H-isochromenes, respectively (Fig. 3b).

Comparison of secondary metabolite gene clusters in the three Monascus species

In order to measure genetic and evolutionary similarity of the enzymes related in these three species, we have analyzed their whole genome sequences and annotated the sequences of the enzymes related in the biosynthesis pathways of these metabolites. Illumina MiSeq 300 bp paired-end sequencing generated 17 million high-quality short reads in individual Monascus species, which corresponds to 200 times the genome size. We can expect that the genome size for each species is around 24 Mb, since in the literature, the genome sizes reported for M. purpureus YY-1 and M. ruber NRRP 1597 were 24.1 Mb [12] and 24.8 Mb (https://genome.jgi.doe.gov/Monru1/Monru1.info.html), respectively. About 94.7% of the short reads satisfied the sequencing error rate (Q20 = 1%). There were a total of 8891 genes identified in M. purpureus, 8645 in M. ruber, and 8647 in M. pilosus. These numbers are similar to the number of genes (7491) identified in M. purpureus YY-1 [12].

Furthermore, to evaluate the clusters of the enzymes related with the secondary metabolic pathways we analyzed their genome sequences using antiSMASH 4.0 [31], and we obtained 24 gene clusters for M. pilosus, 48 for M. ruber, and 20 for M. purpureus. Subsequently, we constructed a dendrogram using the BLAST alignment scores of the 92 gene clusters (Fig. 4), obtaining 22 groups with multiple gene clusters and 32 singletons. Gene clusters with identical gene organization were merged into a single group. The gene-coding regions were predicted using BLASTX search against UniProtKB/Swiss-Prot database (E-value < 1.0 × 10^{-10}).

We also obtained 54 groups of gene clusters (Table 1) that are displayed as Venn diagrams (Fig. 5). Gene clusters were characterized as those for secondary metabolites or ABC transporters based on similarity to known genes [32]. ABC transporters are associated with PKS and NRPS gene clusters in several fungi and are responsible for the export of the corresponding secondary metabolites [33, 34].

Taking gene organization in gene clusters into consideration, only two gene clusters were common in the three species. In all, 85% of the identified secondary metabolic biosynthetic gene clusters were assigned to each Monascus species. Thus, the organization of genes in the genome of each Monascus species is highly diverse (Fig. 5a1). Otherwise, if gene-clusters are classified based on DNA sequence similarity, 16 (30%) were found to be common in the three species. M. ruber has 27 intrinsic gene clusters, while M. pilosus and M. purpureus only have two and three intrinsic gene clusters, respectively (Fig. 5a2), indicating that the number of gene clusters in M. ruber is different from the other two species. Classification for individual gene clusters was conducted using antiSMASH, which identified 8 Type I polyketide biosynthetic systems (T1PKS), 8 non-ribosomal peptide biosynthesis systems (NRPS), three complex systems of T1PKS and NRPS (T1PKS-NRPS), three terpenoid systems (T1PKS-terp), five Type II polyketide biosynthetic systems (T2PKS), and five non-ribosomal peptide biosynthesis systems (NRPS-terp).
biosynthetic systems, 27 putative fatty acid biosynthetic systems, and five others (Fig. 5b). The results of the present study strongly suggest that the three Monascus species have greatly diverged gene clusters; thus, they should be regarded as different species based on chemical taxonomy associated with the production of secondary metabolites.

We further analyzed the differences in the Monascus species at the DNA level by comparing the 8144-bp region where a Monascus azaphilone pigment biosynthetic gene cluster was localized [35]. The sequence analysis was performed for three M. ruber strains (NRRP 1596, JF83291.6, and NBRC 4483), two M. purpureus strains (NRRP 1597 and NBRC 4478), and two M. pilosus strains (NCBI and NBRC 4520), which identified nucleotide differences, including substitutions and insertions/deletions, at 276 bp positions (Table 2). In total, 275 positions (99.6%), except the 5194th nucleotide position, were different between the strains of (i) M. ruber (NRRP 1596 and NBRC 4483) and M. pilosus (NBRC 4520 and NCBI) and strains of (ii) M. purpureus (NRRP 1597 and NBRC 4478) and M. ruber (JF83291.6). It should be noted that NRRP 1597 and NBRC 4483 were assigned to group (i) and JF83291.6 to group (ii). M. pilosus is treated as a synonym of M. ruber in the concatenated phylogeny based on the ITS, BenA, CaM LSU, and RPB2 gene regions of 46 relative strains [36]. However, the evidence obtained in this study indicates that M. pilosus and M. purpureus are different, while M. ruber has similar biosynthetic gene clusters for citrinin, monacolin K, and Monascus azaphilone pigments with M. pilosus and M. purpureus.

**Comparison of citrinin biosynthetic gene clusters in the three Monascus species**

The metabolite analysis using LC-MS shown in Fig. 1b, it was shown that citrinin was produced only by M. purpureus. This raises the question of whether the biosynthetic pathway of citrinin can be regulated negatively via PD culture and whether this pathway exists in M. pilosus and M. ruber genomes. Previous studies have reported that the production of citrinin by M. ruber can be changed by altering the medium and/or culture conditions [37, 38]. To address this issue, we compared the peptide sequences of citrinin biosynthetic genes in the three Monascus species with those in M. purpureus reported by Shimizu et al. [39] and Chen et al. [40]. There are six genes associated with the citrinin biosynthetic pathway, consisting of
Table 1 Classification of gene clusters (ID1–54) corresponding to the dendrogram in Fig. 4

| ID | Type of GCs | Pi | Ru | Pu | Secondary metabolic pathways detected in DNA sequence homology | ATP-binding cassettes | Identity level |
|----|-------------|----|----|----|-------------------------------------------------------------|----------------------|----------------|
| 1  | t1pks       | 1  | 4  | –  | Narbonolide/10-deoxymethanolide (pikAI – AIV), Phthiocerol/phenolphthionocerol (ppsA - E) | –                     | **             |
| 2  | t1pks       | 2  | –  | –  | Phthiocerol/phenolphthionocerol (ppsA - E)                  | –                     | –              |
| 3  | t1pks       | 18 | 19 | 7  | Azaphilone*, Lovastatin (LOVB, LOVF), Narbonolide/10-deoxymethanolide (pikAI – AIV), Phthiocerol/phenolphthionocerol (ppsA - E) | –                     | –              |
| 4  | t1pks       | 22 | 2  | –  | –                                                          | –                     | –              |
| 5  | t1pks       | 23 | 48 | 17 | Lovastatin (LOVB, LOVF), Narbonolide/10-deoxymethanolide (pikAI – AIV), Phthiocerol/phenolphthionocerol (ppsA-C, E) | –                     | *              |
| 6  | t1pks       | 19 | 43 | –  | Monacolin K*, Compactin, Lovastatin (LOVA, LOVB)            | –                     | *              |
| 7  | t1pks       | –  | 8  | –  | Citrinin*, Narbonolide/10-deoxymethanolide (pikAI–AIV), Phthiocerol/phenolphthionocerol (ppsA-C, E) | –                     | *              |
| 8  | t1pks       | 9  | –  | 10 | Byssochlamic acid*, Narbonolide/10-deoxymethanolide (pikAI–AIV), Phthiocerol/phenolphthionocerol (ppsA-C, E) | –                     | *              |
| 9  | nrps        | 4  | 10 | 19 | Neosartorin*, Fengycin (fenA–E), Surfacin (srfAA-AC)        | –                     | **             |
| 10 | nrps        | 13 | 30 | 20 | Brevianamide F (FTMA)                                       | ABCB **              |                |
| 11 | nrps        | 7  | 2  | 8  | Fengycin (fenA-E), Surfacin (srfAA-AC)                      | ABCCC **             |                |
| 12 | nrps        | 11 | 23 | 2  | Fengycin (fenA-D)                                           | ABCCC **             |                |
| 13 | nrps        | 5  | 22 | 12 | Fengycin (fenA-D)                                           | –                     | *              |
| 14 | nrps        | 24 | 21 | 9  | Fengycin (ppsA-D), surafing (srfAA-AC)                      | ABCB **              |                |
| 15 | nrps        | 3  | 11 | –  | Fengycin (fenA-D)                                           | –                     | *              |
| 16 | nrps        | –  | 9  | 8  | Fengycin (fenB, D), Fierrocin (SIDC, SIDD)                  | –                     | –              |
| 17 | t1pks-nrps  | 8  | 40 | 5  | NG-391*, Fengycin (fenA–E), Surfacin (surAA-AC)             | ABCCC **             |                |
| 18 | t1pks-nrps  | 14 | 27 | 14 | Gramicidin (grsa-A), Fengycin (fenA–E), Surfacin (surAA-AC) | –                     | *              |
| 19 | t1pks-nrps  | 20 | 13 | 3  | Lovastatin (LOVB, LOVF), Fengycin (fenA-E), Surfactin (srfAA-AC) | ABCB **              |                |
| 20 | terpene     | 12 | 24 | 11 | Fernesyl-diphosphate (FDFT1)                                | –                     | **             |
| 21 | terpene     | 17 | 12 | 13 | Lupeol (LUP1,2,4,5), Arabidiol (PEN1), Tinuccalladienol (PEN3), seco-amyrin (PEN6) | –                     | *              |
| 22 | terpene     | –  | 15 | –  | –                                                          | –                     | –              |
| 23 | others      | 15 | 42 | 18 | Fengycin (fenA, C, E), Surfacin (srfAA-AB)                  | ABCF **              |                |
| 24 | others      | 6  | 32 | 6  | Kinesin (KIDFC1–3)                                         | –                     | *              |
| 25 | others      | 10 | 5  | 16 | Abscisic aldehyde (AAO1–4), Fengycin (fenA, B, E), Surfactin (srfAA, AB) | ABCB **              |                |
| 26 | others      | 16 | 33 | 4  | Histidinol (hisD, IE)                                      | –                     | *              |
| 27 | others      | 21 | 28 | 1  | –                                                          | –                     | *              |
| 28 | cf-putative | –  | 1  | –  | –                                                          | –                     | –              |
| 29 | cf-putative | 3  | 1  | –  | Palmitin (ZDHHC)                                           | –                     | –              |
| 30 | cf-putative | 6  | 1  | –  | Palmitin (ZDHHC4)                                          | –                     | –              |
| 31 | cf-putative | –  | 7  | –  | –                                                          | –                     | –              |
| 32 | cf-putative | 8  | 1  | –  | –                                                          | –                     | –              |
| 33 | cf-putative | 9  | 1  | –  | –                                                          | –                     | –              |
| 34 | cf-putative | 11 | 1  | –  | –                                                          | –                     | –              |
| 35 | cf-putative | 14 | 1  | –  | Serine, Threonine (PP1C, 2C, 4C, 6C)                       | –                     | –              |
| 36 | cf-putative | 15 | 1  | –  | –                                                          | –                     | –              |
| 37 | cf-putative | 16 | 1  | –  | –                                                          | –                     | –              |
| 38 | cf-putative | 17 | 1  | –  | –                                                          | –                     | –              |
| 39 | cf-putative | 18 | 1  | –  | Mannan (ANPI, MMN9)                                        | –                     | –              |
Table 1  Classification of gene clusters (ID1–54) corresponding to the dendrogram in Fig. 4 (Continued)

| ID | Type of GCs | Pi | Ru | Pu | Secondary metabolic pathways detected in DNA sequence homology | ATP-binding cassettes | Identity level |
|----|-------------|----|----|----|-------------------------------------------------------------|----------------------|----------------|
| 40 | cf-putative  | 20 | –  | –  | –                                                            | –                    | –              |
| 41 | cf-putative  | 25 | –  | –  | –                                                            | –                    | –              |
| 42 | cf-putative  | 29 | –  | –  | –                                                            | –                    | –              |
| 43 | cf-putative  | 31 | –  | –  | –                                                            | –                    | –              |
| 44 | cf-putative  | 34 | –  | –  | –                                                            | –                    | –              |
| 45 | cf-putative  | 36 | –  | –  | –                                                            | –                    | –              |
| 46 | cf-putative  | 37 | –  | –  | –                                                            | –                    | –              |
| 47 | cf-putative  | 38 | –  | –  | –                                                            | –                    | –              |
| 48 | cf-putative  | 39 | –  | –  | Lovastatin (LOVB-G), Phthiocerol/Phenolphthiocerol (ppsA,C) | –                    | –              |
| 49 | cf-putative  | 44 | –  | –  | –                                                            | –                    | –              |
| 50 | cf-putative  | 45 | –  | –  | –                                                            | –                    | –              |
| 51 | cf-putative  | 46 | –  | –  | –                                                            | –                    | –              |
| 52 | cf-putative  | 47 | –  | –  | –                                                            | –                    | –              |
| 53 | cf_fatty_acid | 26 | –  | –  | Fatty acid (FAS1,2)                                         | –                    | –              |
| 54 | cf_fatty_acid | 35 | –  | –  | Fatty acid (FAS1,2)                                         | –                    | –              |

The type of GCs was determined using antiSMASH software. Pi, Ru, and Pu represent the cluster ID in Fig. 4. Identical gene organization is denoted by red numbers. The secondary metabolic pathways represent the secondary metabolite information based on DNA sequence homology. Type of ATP-binding cassettes detected in individual groups is represented as ATP-binding cassettes. Gene-cluster groups with both identical gene organization and high DNA sequence similarity are denoted by "**" and groups with only high DNA sequence similarity are denoted by "*". Two-letter abbreviations use used for the Venn diagrams: Pi, M. pilosus NBRC 4520; Ru, M. ruber NBRC 4483; Pu, M. purpureus NBRC 4478.

Fig. 5  Venn diagrams of secondary metabolite biosynthetic gene clusters observed in three Monascus species. (a1) Venn diagram classifying 82 secondary metabolic synthesis gene clusters of three Monascus species. (a2) Venn diagram classifying 82 secondary metabolic synthesis gene clusters of three Monascus species based on DNA sequence homology. The total number of gene clusters was 54. Venn diagram classifying 54 secondary metabolic synthesis gene clusters of three Monascus species based on DNA sequence homology: T1PKS (b1), NRPS (b2), T1-PKS-NRPS (b3), Terpenes (b4), Others (b5), Cf-Putative or Cf-fatty acid (b6).
*pk*/*CT* (encoding polyketide synthase, *PksCT/CitS*), *mpl*/*1* (a serine hydrolase, *CitA*), *mpl*/*2* (an iron II oxidase, *CitB*), *mpl*/*7* (an oxidoreductase, *CitC*), *mpl*/*4* (encoding an aldehyde dehydrogenase, *CitD*), and *mpl*/*6* (a short chain dehydrogenase, *CitE*). Table 3 shows the homologous regions of individual genes that aligned with the reference sequences (E-value < $10^{-44}$). Two proteins, *CitB* and *CitS*, were shorter in *M. ruber* and *M. pilosus* than the reference sequences. However, the same proteins were conserved in *M. purpureus* compared with the reference sequences. It suggests that the citrinin biosynthetic genes, *mpl*/*2* (*CitB*), and *pk*/*CT* (*CitS*) were incomplete in *M. pilosus* NBRC 4520 and *M. ruber* NBRC 4483, and consequently, citrinin production is blocked in these species.

### Comparison of monacolin biosynthetic gene clusters in the three *Monascus* species

Monacolin K was first isolated from the medium of *M. ruber* [41] and its biosynthesis pathway was determined, in *M. pilosus* (BCRC 387072), composed of nine enzymes that have a high level of homology with genes in the monacolin K biosynthetic gene cluster of *Aspergillus terreus* [42]. Three strains of *M. purpureus*, specifically NRRP 1596, YY-1 (an industrial strain), and KACC 42430 (a laboratory strain), lack an intact monacolin K gene cluster [43]. In the present study, we also examined the monacolin gene clusters in three *Monascus* species. As shown in Table 4, all 9 Mok genes, especially MokC, MokD, and MokF, were shorter in *M. purpureus* NBRC 4478 than *M. pilosus* NBRC 4520 and *M. ruber* NBRC 4483. Thus, *M. purpureus* NBRC 4478 also lacks the complete monacolin K gene sequence.

### Discussion

The three *Monascus* species examined in the present study are commonly used for food fermentation in the cuisine of East Asian countries [1–3]. Citrinin, a nephrotoxic agent, was reportedly produced in *M. purpureus* but not in *M. pilosus* [30, 44, 45]. This is corroborated by the present results from the metabolome and genome analyses revealing that citrinin biosynthetic pathways in *M. pilosus* were incomplete compared with those from *M. purpureus*. The three *Monascus* species can produce ubiquitous and species-specific pigment-related compounds (Figs. 2 and 3). Analysis of gene-organization revealed 54 greatly diverged gene clusters in the three *Monascus* species studied (Fig. 5a2). Furthermore, comparison of a 8144 bp region, in which a gene cluster of *Monascus* azaphilone synthases was localized, revealed that *M. pilosus* and *M. purpureus* can be clearly distinguished at the nucleotide level. In addition, *M. ruber* NBRC 4483 and NRRP 1597 have highly similar DNA sequences with *M. pilosus*; however, *M. ruber* JF83291.6 has highly similar DNA sequences with *M. purpureus* (Table 2). Though in some phylogenetic studies [36, 46] *M. pilosus* and *M. ruber* were not distinguished, in our analysis, their phenotypes distance clearly distinct. On the other hand, taking the pigment biosynthetic gene clusters into consideration, *M. pilosus* and *M. purpureus* should be defined as different groups. Thus, based on the findings of the present study, the *Monascus* species studied here can be classified into two groups: (i) the *M. pilosus* clade and (ii) the *M. purpureus* clade. And the results shown in Table 2 suggests that there may some *M. ruber* strains which can be related with each clade.

The mycotoxin citrinin is produced by various *Penicillium*, *Aspergillus*, and *Monascus* species [40, 44, 45]. Previously studied *M. purpureus* strains (ATCC 16365 in Java, 16379 in Taiwan, IFO 30873, and DSM 1379 by [40, 47]; YY1 by [14] can produce citrinin as a secondary metabolite. However, among the *Monascus* species, two *M. pilosus* strains (BCRC 38072 in Taiwan by [40]; NBRC 4520 in this study) cannot produce citrinin. Interestingly, several previously reported *M. ruber* strains, particularly ATCC 16246, 16378, 16366, 18199, 16371, and 18199 by Chen et al. [40], AUMC 4066 (CBS109.07) and AUMC 5705 by Moharram et al. [48], NRRP 1597 by Kwon [43], and NBRC 4483 in this study, lack citrinin production activities, but other strains, such as Tiegh by Ostry et al. [47] and ATCC 96218 by Hajjaj et al. [38] have the potential to produce this secondary metabolite. Thus, *M. ruber* can be classified into citrinin-producing and non-citrinin producing types. Based on the comparison of citrinin biosynthetic proteins, the former type might correspond to *M. purpureus* strains and the latter to *M. pilosus* strains.

In the analysis of the monacolin K gene cluster, four *M. purpureus* strains, specifically NRRP 1596, YY-1, KACC 42430 [43], and NBRC 4478 (in this study), lack an intact monacolin K gene cluster. By contrast, *M. pilosus* NBRC 4520 and *M. ruber* NBRC 4483 have a complete set of monacolin K gene clusters. Thus, it should be noted that *M. pilosus* NBRC 4520 and *M. ruber* NBRC 4483 can produce monacolin K but lack a complete set of citrinin biosynthetic gene clusters.

The classification of strains according to the two clade groups designated as (i) *M. pilosus* and (ii) *M. purpureus* may play an important role in the food industry and industrial field through the improved utilization of *Monascus* species. However, in view of food safety, we need to confirm whether the toxins produced by some *Monascus* strains exist in the genome or metabolome. Metabolites are generally classified into primary metabolites that are essential for growth and reproduction and secondary metabolites that are usually involved in mechanisms for ecological adaptation but are not essential to regular cellular processes. Metabolic pathways can be divided into
Table 2 Comparison of six *M. purpureus*, *M. ruber*, and *M. pilosus* strains based on Monascus azaphilone pigment polyketide synthase gene

| M. pilosus 4520 | A | A | A | T | T | G | G | G | A | T | T | A | G | C | G | A | T | A | A | A | A | A | T | C |
| M. pilosus 4520 | A | A | A | T | T | G | G | G | A | T | T | A | G | C | G | A | T | A | A | A | A | A | T | C |
| M. ruber NRLP | A | A | A | T | T | G | G | G | A | T | T | A | G | C | G | A | T | A | A | A | A | A | T | C |
| M. ruber 4483 | A | A | A | T | T | G | G | G | A | T | T | A | G | C | G | A | T | A | A | A | A | A | T | C |
| M. ruber JF83291.6 | C | G | G | C | C | C | A | A | G | C | C | G | A | G | A | G | C | G | G | G | C | G | C | T |
| M. purpureus v1_0 | C | G | G | C | C | C | A | A | G | C | C | G | A | G | A | G | C | G | G | G | C | G | C | T |
| M. purpureus 4478 | C | G | G | C | C | C | A | A | G | C | C | G | A | A | A | A | A | A | A | T | C | G | C | C |
| M. pilosus 4520 | T | A | G | A | T | T | T | T | A | A | G | A | A | G | G | G | T | T | T | A | T | A | T | C | G | C | C |
| M. pilosus 4520 | T | A | G | A | T | T | T | A | A | G | A | A | G | G | G | T | T | T | A | T | A | T | C | G | C | C |
| M. ruber NRLP | T | A | G | A | T | T | T | A | A | G | A | A | G | G | G | T | T | T | A | T | A | T | C | G | C | C |
| M. ruber JF83291.6 | C | G | G | C | C | C | A | A | G | C | C | G | A | A | C | C | G | A | C | C | T | A | T | T |
| M. purpureus v1_0 | C | G | G | C | C | C | A | A | G | C | C | G | A | A | C | C | G | A | C | C | T | A | T | T |
| M. purpureus 4478 | C | G | G | C | C | C | A | A | G | C | C | G | A | A | A | A | C | C | G | A | C | C | T | A | T |
| M. pilosus 4520 | T | A | G | A | T | T | T | T | T | T | T | T | T | T | T | G | A | A | A | A | G | A | A | - | - | - | A |
| M. pilosus 4520 | T | A | G | A | T | T | T | T | T | T | T | T | T | T | T | G | A | A | A | A | G | A | A | - | - | - | A |
| M. ruber NRLP | C | T | C | T | T | A | A | T | T | T | T | T | T | T | T | G | A | A | A | A | G | A | A | - | - | - | A |
| M. ruber JF83291.6 | T | C | T | C | C | G | G | C | C | C | C | C | C | C | C | C | C | T | T | G | C | A | C | G | A | C | G |
| M. purpureus v1_0 | T | C | T | C | C | G | G | C | C | C | C | C | C | C | C | C | T | T | G | C | A | C | G | A | C | G |
| M. purpureus 4478 | T | C | T | C | C | G | G | C | C | C | C | C | C | C | C | C | C | T | T | G | C | A | C | G | A | C | G |
| M. pilosus 4520 | T | A | A | C | T | C | A | A | T | G | T | A | A | T | A | T | C | T | A | A | A | T | T | G | T |
| M. pilosus 4520 | T | A | A | C | T | C | A | A | T | G | T | A | A | T | A | T | C | T | A | A | A | T | T | G | T |
| M. ruber NRLP | T | A | A | C | T | C | A | A | T | G | T | A | A | T | A | T | C | T | A | A | T | T | G | T |
| M. ruber JF83291.6 | C | G | G | T | C | G | G | T | C | C | C | C | C | T | G | C | T | C | T | C | C | C | C | C | C | A | G |
| M. purpureus v1_0 | C | G | G | T | C | G | G | T | C | C | C | C | C | C | C | C | T | C | T | C | C | C | C | C | C | A | G |
| M. purpureus 4478 | C | G | G | T | C | G | G | T | C | C | C | C | C | C | C | C | T | C | T | C | C | C | C | C | C | A | G |

Higa et al. BMC Genomics (2020) 21:679

Page 10 of 17
Table 2 Comparison of six *M. purpureus*, *M. ruber*, and *M. pilosus* strains based on Monascus azaphilone pigment polyketide synthase gene (Continued)

| Strain            | 7189 | 7191 | 7241 | 7270 | 7316 | 7325 | 7397 | 7415 | 7421 | 7445 | 7457 | 7556 | 7583 | 7631 | 7676 | 7692 | 7769 | 7853 | 7874 | 7976 | 8043 | 8044 | 8083 |
|-------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| *M. pilosus* 4520 | C    | A    | T    | A    | C    | T    | C    | C    | G    | G    | G    | C    | A    | T    | C    | C    | A    | G    | T    | A    | T    | T    |
| *M. pilosus* 4520 | C    | A    | T    | A    | C    | T    | A    | T    | C    | C    | G    | G    | C    | A    | T    | C    | C    | A    | G    | T    | A    | T    |
| *M. ruber* NRLP  | C    | A    | T    | A    | C    | T    | A    | T    | C    | C    | G    | G    | C    | A    | T    | C    | C    | A    | G    | T    | A    | T    |
| *M. ruber* 4483  | C    | A    | T    | A    | C    | T    | A    | T    | C    | C    | G    | G    | C    | A    | T    | C    | C    | A    | G    | T    | A    | T    |
| *M. purpureus* v1_0 | T    | G    | G    | G    | T    | C    | G    | C    | G    | T    | A    | C    | G    | G    | A    | T    | T    | G    | A    | G    | G    | C    | C    |
| *M. purpureus* 4478 | T    | G    | G    | G    | T    | C    | G    | C    | G    | T    | A    | C    | G    | G    | A    | T    | T    | G    | A    | G    | G    | C    | C    |
| *M. pilosus* 4520 | A    | A    | G    | A    | C    | T    | C    | A    | T    | C    | T    | A    | T    | A    | G    | G    | A    | C    | C    | T    | T    | G    | T    |
| *M. pilosus* 4520 | A    | A    | G    | A    | C    | T    | C    | A    | T    | C    | T    | A    | T    | A    | G    | G    | A    | C    | C    | T    | T    | G    | T    |
| *M. ruber* NRLP  | A    | A    | G    | A    | C    | T    | C    | A    | T    | C    | T    | A    | T    | A    | G    | G    | A    | C    | C    | T    | T    | G    | T    |
| *M. ruber* 4483  | A    | A    | G    | A    | C    | T    | C    | A    | T    | C    | T    | A    | T    | A    | G    | G    | A    | C    | C    | T    | T    | G    | T    |
| *M. ruber* JF83291.6 | G    | C    | C    | G    | T    | C    | T    | C    | C    | T    | C    | C    | G    | C    | G    | A    | A    | G    | T    | G    | C    | G    | C    |
| *M. purpureus* v1_0 | G    | C    | C    | G    | T    | C    | T    | C    | C    | T    | C    | C    | G    | C    | G    | A    | A    | G    | T    | G    | C    | G    | C    |
| *M. purpureus* 4478 | G    | C    | C    | G    | T    | C    | T    | C    | C    | T    | C    | C    | G    | C    | G    | A    | A    | G    | T    | G    | C    | G    | C    |
| Strain          | Position          | Sequence                                                                 |
|-----------------|-------------------|--------------------------------------------------------------------------|
| M. purpureus v1_0 | 4478              | C G C C A C T G G C C A C A G G G A T G C C C T G                       |
| M. ruber JF83291.6 | 1601              | T T T T A A A C C G G G A G T C A G T T T T T C T C G T G               |
| M. pilosus 4520  | 4 4 8 3           | A T T A A A C C G G G A G T C A G T C C T G T G                         |

Continued...
| Strain          | Sequence | 8105 | 8107 |
|----------------|----------|------|------|
| *M. pilosus* 4520 | T A T A C T T T T A A T C C G T G T T T A C A T T G C |      |      |
| *M. pilosus* 4520 | T A T A C T T T T A A T C C G T G T T T A C A T T G C |      |      |
| *M. ruber* NRLP  | T A T A C T T T T A A T C C G T G T T T A C A T T G C |      |      |
| *M. ruber* 4483  | T A T A C T T T T A A T C C G T G T T T A C A T T G C |      |      |
| *M. ruber* JF83291.6 | A G C G T C C C C G C T T C C C C C C G - - - C A T |      |      |
| *M. purpureus* v1_0 | A G C G T C C C C G C T T C C C C C C G - - - C A T |      |      |
| *M. purpureus* 4478 | A G C G T C C C C G C T T C C C C C C G - - - C A T |      |      |

Table 2: Comparison of six *M. purpureus*, *M. ruber*, and *M. pilosus* strains based on Monascus azaphilone pigment polyketide synthase gene (Continued)
two types: one is the general pathway shared by most fungi and the other are specialized pathways that have evolved in response to specific ecologies of certain lineages and are consequently more narrowly distributed at the taxonomic level. The citrinin pathway belongs to the former as it is present in many *Penicillium*, *Aspergillus*, and *Monascus* species [30, 44, 45]. However, the biosynthetic gene cluster of *Monascus* azaphilone pigments is limited in the *Monascus* genera. The biosynthetic process of secondary metabolites forms a cluster or non-clustered gene organization that is integral to the entire spectrum of fungal ecological strategies (e.g., saprotrophic, pathogenic, and symbiotic). Gene duplication (GD) is often implicated in the evolution of fungal metabolism (Floudas et al., 2012). A second source of metabolic gene innovation in fungi is horizontal gene transfer (HGT), which includes xenobiotic catabolism [49], toxin production [50], and degradation of plant cell walls [51]. GD and HGT were more frequently found in clustered genes than in their non-clustered counterparts [52]. In the biosynthetic gene clusters of *Monascus* azaphilone pigments and citrinin, the common trends in the strains of the three *Monascus* species are explained by the suggested *M. pilosus* and *M. purpureus* clades, whereas *M. ruber* has either *M. pilosus* or *M. purpureus* characteristics. *Monascus*-specific diverged pigments may have evolved because of GD and HGT events, resulting in the creation of clustered genes in their genomes; thus, a large number of gene clusters was observed (Table 1). Chemotaxonomy, including pigment production, is the most useful way to study the divergence of *Monascus* genera. Here, we compared the PKS responsible for the biosynthesis of azaphilone pigment from three *Monascus* species (*M. pilosus*, *M. purpureus* and *M. ruber*) and six strains. More genome sequences of *Monascus* species will need to be determined to better understand the production of secondary metabolites in these organisms.

### Conclusions

In this study, the complete genome sequences of *M. pilosus* NBRC 4520, *M. purpureus* NBRC 4478, and *M. ruber* NBRC 4483 were obtained. Three biosynthetic gene clusters, specifically monacolin K, citrinin, and

| Protein | *M. pilosus* NBRC 4520 | *M. ruber* NBRC 4483 | *M. purpureus* NBRC 4478 | Reference peptide sequence |
|---------|------------------------|----------------------|--------------------------|---------------------------|
| CitA (mpl1) | 249 (79.6) | 249 (79.6) | 312 (99.7) | 313 |
| CitB (mpl2) | 229 (69.6) | 229 (69.6) | 328 (99.7) | 329 |
| CitD (mpl4) | 458 (91.4) | 458 (91.4) | 500 (99.8) | 501 |
| CitE (mpl6) | 231 (79.1) | 231 (79.1) | 284 (97.3) | 292 |
| CitC (mpl7) | 532 (85.5) | 532 (85.5) | 618 (99.4) | 622 |
| CitS (pksCT) | 1525 (58.8) | 1525 (58.8) | 2396 (92.4) | 2593 |

Reference peptide sequence is *Monascus purpureus* citrinin biosynthesis gene clusters (https://www.ncbi.nlm.nih.gov/nuccore/AB243687.1)

| Protein | *M. pilosus* NBRC 4520 | *M. ruber* NBRC 4483 | *M. purpureus* NBRC 4478 | Reference peptide sequence |
|---------|------------------------|----------------------|--------------------------|---------------------------|
| MokA | 1910 (100) | 1910 (100) | 1772 (92.8) | 1910 |
| MokB | 1947 (100) | 1947 (100) | 1501 (77.1) | 1947 |
| MokC | 339 (100) | 339 (100) | 28 (8.3) | 339 |
| MokD | 263 (100) | 263 (100) | 154 (58.6) | 263 |
| MokE | 265 (100) | 265 (100) | 231 (87.2) | 265 |
| MokF | 258 (100) | 258 (100) | 167 (64.7) | 258 |
| MokG | 996 (99.0) | 977 (100) | 322 (32.3) | 977 |
| MokH | 487 (100) | 487 (100) | 42 (8.6) | 487 |
| MokI | 174 (100) | 174 (100) | 107 (61.5) | 174 |

Reference peptide sequence is *Monascus pilosus* monacolin biosynthesis gene clusters (https://www.ncbi.nlm.nih.gov/nuccore/DQ176595.1)
azaphilone pigments that are involved in secondary metabolism, were analyzed and compared. The grouping of strains according to the two clade groups, designated as (i) M. pilosus and (ii) M. purpureus, may play an important role in the food industry and industrial field through the improved utilization of Monascus species. The PKS genes responsible for the biosynthesis of azaphilone pigment from the three species were compared. This genome-based analysis showed M. ruber could not be clearly grouped as a species. However, in view of food safety, further studies are needed to confirm whether the toxins produced by some Monascus strains originate from the genome and not from the metabolome.

Methods

Strains, culture conditions, and metabolite detection

Three Monascus species, specifically, M. pilosus NBRC 4520, M. purpureus NBRC 4478, and M. ruber NBRC 4483, were obtained from the National Institute of Technology and Evaluation in Japan. The three species were cultured in potato dextrose liquid medium at 30 °C for 7 days with 140 rpm shaking in TAITC BR-23FP. A solution of 10 mg freeze-dried PDL medium added with 1 mL methanol was sonicated for 30 min to extract secondary metabolites. The extracted metabolites were measured using a Shimadzu LCMS-8040 system (Shimadzu, Kyoto, Japan) with 300 mm ODS MonoBis columns (Kyoto Monotech Co., Ltd., Kyoto, Japan). Each metabolite was estimated based on the m/z of each peak, referring to the m/z of the metabolites previously reported in Monascus. We measured three replicates for each species and applied two-dimensional hierarchical clustering to visualize their similarity, using the Euclidian distance of their profiles of the observed pigments concentration as a measure of similarity score and applying the Ward method.

Genome sequencing and assembly

We isolated genomic DNA from the three species individually and sequenced them using Illumina MiSeq paired-end libraries (300 bp read each end and 500 inserts). Approximately 8.5 million reads (around 5 Gb) for each sample were obtained and assembled using ABySS 2.0 de novo assembler [28]. We obtained 5000 to 10,000 assembled scaffolds for each samples and the N50 value of the total scaffolds were 133 Kb. The accumulated total length of the assembled contigs was 24.8 M bp, which is close to that of M. purpureus NRRP 1596 genome (ATCC 16365) with 23.4 Mb [12] and M. ruber NRRP 1597 (ATCC 13692) with 24.9 Mb [12]. The raw read sequences and the assembled contig sequences are deposited to the DNA Data Bank of Japan (DDBJ) and available under accession numbers DRX224643, DRX224644, DRX224645, respectively. To identify the gene-coding regions, the nucleotide sequence of the assembled scaffolds was annotated using DIAMOND, a high throughput BLASTX compatible sequence alignment algorithm [53]. The assembled sequences were also analyzed by BLASTed using the library of the whole UniProtKB/Swiss-Prot database [54]. Annotated genes of M. purpureus NRRP 1596 and M. ruber NRRP 1597 for were used for validation, with a cutoff E-value <1E-10. We further analyzed the genomes using antiSMASH pipeline [55] to extract the functional gene clusters such as PKS, in each Monascus species.

Acknowledgements

The authors would like to thank AXIOHELIX Co. Ltd. for their support regarding genome acquisition.

Authors’ contributions

Conceptualization and design of the study were performed by YH, SK, and NO. Sample preparation and genomic DNA isolation were carried out by YH. Assembly and scaffolding of sequencing reads were performed by NO. Subsequent comparative genomic analysis were conducted by NO and YH. Statistical processing and figure creation were conducted by SK. Culture and LC-MS analysis were performed by YH. Valuable comments and advice on presenting the results were provided by AA, MK, and YSK. All authors have read and approved the final manuscript.

Funding

This work was supported by the Ministry of Education, Culture, Sports, Science, and Technology of Japan (16 K07223 and 17 K00406), Platform Project for Supporting Drug Discovery and Life Science Research funded by the Japan Agency for Medical Research and Development (18 am0101111), the National Bioscience Database Center (NBDC) and NAIST Bigdata Project.

Availability of data and materials

Raw sequence and assembled contig data are available at the DDBJ Sequence Read Archive (https://www.ddbj.nig.ac.jp/dra/) under accession numbers DRX224643, DRX224644, DRX224645.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

Not applicable.

Author details

1R&D Center, Kobayashi Pharmaceutical Co., Ltd, Ibaraki-shi, Toyokawa, 1-30-3, Osaka, Japan. 2Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma-shi, Takayama-cho, Nara 8916-5, Japan. 3Data Science Center, Nara Institute of Science and Technology, Ikoma-shi, Takayama-cho, Nara 8916-5, Japan.

Received: 22 January 2020 Accepted: 23 June 2020

References

1. Lee CL, Pan TM. Development of Monascus fermentation technology for high hypolipidemic effect. Appl Microbiol Biotechnol. 2012;94:1449–59.
2. Wang TH, Lin TF. Monascus rice products. Adv Food Nutr Res. 2007;53:123–59.
3. Yasuda M, Tachibana S, Kuba-Miyara M. Biochemical aspects of red koji and tofu prepared using Monascus fungi. Appl Microbiol Biotechnol. 2012;96:49–60.
4. Gao JM, Yang SX, Qin JC. Azaphilones: chemistry and biology. Chem Rev. 2013;113:4755–81.
5. Mostafa ME, Abbady MS. Secondary Metabolites and Bioactivity of the Monascus Pigments Review Article. Biotechnol Biochem. 2014;89(1):1–13.

6. Pastrana L, Blanc PJ, Santerre AL, Loret MO, Goma G. Production of red pigments by Monascus ruber in synthetic media with a strictly controlled nitrogen source. Process Biochem. 1995;30:333–41.

7. Wanping C, Feng Y, Molnár I, Chen F. Nature and nurture: Confluence of pathway determinism with metabolic and chemical serendipity diversifies Monascus azaphilone pigments. Nat Prod Rep. 2019;36(4):561–72.

8. Mapari SAS, Meyer AS, Thrane U, Frisvad JC. Identification of potentially safe promising fungal cell factories for the production of polyketide natural food colorants using chemometrically rationalized. Microb Cell Factories. 2009;8:24.

9. Juslin L, Frisvad J, Neijanti P, Nosrati AC, Bayat M, Azar OL. An investigation on measurement means of citrinin toxicity and microbial safety: a systematic review and meta-analysis. Atherosclerosis. 2014;245:75–81.

10. Yu X, Wu H, Zhang J. Effect of Monascus as a nitrite substitute on color, lipid oxidation, and proteolysis of fermented meat mince. Food Sci Technol. 2015;24:257–81.

11. Yong Y, Liu B, Du X, Li P, Liang B, Cheng X, Du L, Huang D, Wang L, Wang S. Complete genome sequence and transcriptomics analyses reveal pigment biosynthesis and regulatory mechanisms in an industrial strain, Monascus purpureus YY-1. Sci Rep. 2015;5:8331–9.

12. Nejati P, Nosrati AC, Bayat M, Azar OL. An investigation on measurement means of citrinin toxicity and microbial safety: a systematic review and meta-analysis. Atherosclerosis. 2015;240:415–23.

13. Carvalho JC, Pandey A, Babitha S, Soccol CR. Production of pigments by Monascus purpureus and Transcriptome Analysis of Genes Involved in Monacolin K Biosynthesis. PLoS One. 2017;12(1):e0170149.

14. Liang B, Du XX, Li P, Sun CC, Wang S. Investigation of citrinin and pigment biosynthesis mechanisms in Monascus purpureus by transcriptomic analysis. Front Microbiol. 2019;10:1371–11.

15. Canvalho IC, Pandey A, Babitha S, Soccol CR. Production of Monascus biopigments: an overview. Agro Food Ind Hi Tech. 2003;14:37–42.

16. Nejati P, Nosrati AC, Bayat M, Azar OL. An investigation on measurement means of citrinin toxicity and microbial safety: a systematic review and meta-analysis. Atherosclerosis. 2015;240:415–23.

17. Chen W, Chen R, Liu Q, He T, He K, Ding X, Kang L, Guo X, Xie N, Zhou Y, Lu Y, Cai R, Mokhar L, Li M, Miao Y, Peng C, Orange, red yellow: biosynthesis of azaphilone pigments in Monascus fungi. Chem Sci. 2017:8:4917–25.

18. Wang Q, Cai H, Yoon YM, Chen CC, Kwon HJ. Genetic localization and in vivo characterization of a Monascus azaphilone pigment biosynthetic gene cluster. Appl Microbiol Biotechnol. 2013;97:6337–45.

19. Barbosa RNA, Leong SL, Vinnere-Pettersson O, Souza-Motta CM, Frisvad JC, Samson RA, Oliver NT, Houbren J. Phylogenetic analysis of Monascus and new species from honey, pollen and nests of stingless bees. Stud Mycol. 2017;86:29–51.

20. Moussa LA, Abdel Azeiz AZ. Effect of media composition on citrinin and bio-pigments production by Monascus ruber. J Appl Biol Biotechnol. 2017;5(1):104–9.

21. Hsu YW, Hsu LC, Liang YH, Kuo YH, Pan TM. Monaphilones A-C, three new monacolin K but citrinin. Food Sci Biotechnol. 2016;25:1115–21.

22. Lau SKP, Yuen KYY. The biosynthetic pathway for a thousand-year-old Monascus ruber pigment biosynthesis. PLoS One. 2017;12(1):e0170149.

23. Juzlova P, Martinkova L, Kren V. Secondary metabolites of the fungus Monascus pilosus in synthetic media with a strictly controlled nitrogen source. Process Biochem. 1995;30:333–41.

24. Floudas D, Sinder M, Bley R, Barry K, Blanchette RA, et al. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Sci. 2008;336:1715–9.

25. Jackson SD, Vandervalk BP, Mohammadi H, Chu J, Yeo S, Hammond SA, Jahesh G, Khan H, Coombe L, Warren RL, Birlo I. ABYS 2.0 resource-efficient assembly of large genomes using a bloom filter. Genome Res. 2017;27:768–77.

26. Rasheva TV, Nedeva TS, Hallet JN, Kujumdzieva AV. Characterization of a new species from honey, pollen and nests of stingless bees. Stud Mycol. 2017;86:29–51.

27. Ostry V, Malir F, Ruprich J. Producers and important dietary sources of ochratoxin a and citrinin. Toxin. 2013;5:1574–86.

28. Moharram AAA, Mostafa ME, Ismail MA. Chemical profile of Monascus ruber strains. Food Technol Biotechnol. 2012;30:490–3.

29. Gardner DM, McDonald MC, Covarelli L, Solomon PS, Rusu AG, et al. Comparative pathogenomics reveals horizontally acquired novel virulence genes in fungi infecting cereal crops. PLoS Pathog. 2012;8:e1002952.

30. Friesen TL, Stuenkenbrock EH, Liu Z, Meinhardt S, Ling H, et al. Emergence of new species as a result of interspecific virulence gene transfer. Nat Genet. 2006;38:953–6.

31. Garcia-Valles S, Romeu A, Palau J. Horizontal gene transfer of glycosyl hydrolases of the rumen fungi. Mol Biol Evol. 2000;17:352–461.

32. Witsecaher JV, Slot JC, Rokas A. The evolution of fungal metabolic pathways. PLoS One. 2014;10:e1004816.1–11.
53. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015; https://doi.org/10.1038/nmeth.3176.

54. Pundir S, Martin MJ, O'Donovan C. UniProt protein knowledgebase. In: Wu C, Arighi C, Ross K, editors. Protein bioinformatics. Methods in molecular biology, vol. 1558. New York: Humana Press; 2017.

55. Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res. 2011; 39(suppl_2):W339–46.

**Publisher's Note**

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