J Appl Biol Chem (2016) 59(1), 45–47
http://dx.doi.org/10.3839/jabc.2016.009

Online ISSN 2234-7941
Print ISSN 1976-0442

Short Communication: Biochemistry/Molecular Biology

Some *Monascus purpureus* Genomes Lack the Monacolin K Biosynthesis Locus

Hyung-Jin Kwon* · Bijinu Balakrishnan · Yeon-Ki Kim

Received: 30 October 2015 / Accepted: 26 November 2015 / Published Online: 31 March 2016

Abstract Two *Monascus purpureus* genomes lack the monacolin K biosynthesis locus (*mok*), while *Monascus* species are generally assumed to be monacolin K producers. These *M. purpureus* harbor a fusion of *mokA* and *mokB* orthologues. This finding suggests that an ancestral *mok* locus underwent a deletion event in the *M. purpureus* genome.

Keywords genome sequence · monacolin K biosynthetic gene cluster · *Monascus purpureus* · polyketide synthase gene

Monascus species (a red yeast) such as *M. pilosus*, *M. purpureus*, and *M. ruber* have been used for food fermentation in eastern Asia, especially in China (Wang and Lin, 2007; Lee and Pan, 2012; Yasuda et al., 2012). A popular *Monascus*-fermented product is red yeast rice, and its health benefit is largely attributed to the presence of monacolin K (MON-K; lovastatin), the well-known medicinal ingredient used to treat hypercholesterolemia (Lee and Pan, 2012). *Monascus* species are also sources for natural food colorants, and its active ingredients are *Monascus* azaphilone pigments (MAzPs) (Feng et al., 2012; Patakova, 2013). MON-K and MAzP are polyketide compounds, and their biosynthetic gene clusters have been identified and characterized. The gene cluster for MON-K has been reported for *M. pilosus* (Chen et al., 2008b), and its identity was confirmed by virtue of its high similarity to that of lovastatin from *Aspergillus terreus* (Kennedy et al., 1999).

H.-J. Kwon · B. Balakrishnan · Y.-K. Kim
Department of Biological Science and Bioinformatics, Myongji University, Yongin 449-728, Republic of Korea

*Corresponding author (H.-J. Kwon: hjink@mju.ac.kr)

The biosynthetic gene cluster for MAzP has been reported in *M. pilosus* and *M. purpureus* (Balakrishnan et al., 2013; 2014), while a MAzP oxidoreductase gene (*MpigE*, the ortholog of *mppC* in *M. purpureus* and *M. pilosus*) has been reported in *M. ruber* (Liu et al., 2014). A biosynthetic gene cluster was also identified for mycotoxin citrinin, which is well known as a *Monascus* polyketide metabolite (Shimizu et al., 2007), and defined by an ectopic gene expression (Sakai et al., 2008).

A DNA sequence-based approach first evaluated species-dependent potential of citrinin biosynthesis among diverse *Monascus* species (Chen et al., 2008a). This report substantiated the absence of citrinin biosynthetic gene(s) in several *Monascus* species, including *M. pilosus* and *M. ruber*, indicating that these strains are intrinsically incapable of producing citrinin. Here, the known polyketide biosynthetic gene clusters were localized in the genome sequences of *M. purpureus* NRRL 1596 (ATCC 16365, CBS 109.07; the type strain of *M. purpureus* Went) and *M. ruber* NRRL 1597 (ATCC 13692) in the genome portal of the U.S. Department of Energy Joint Genome Institute (DOE-JGI). These data can be found at the web pages are http://genome.jgi.doe.gov/Monpu1/Monpu1.home.html and http://genome.jgi.doe.gov/Monru1/Monru1.home.html). First, we examined the *M. ruber* genome for a citrinin biosynthetic gene cluster. Previous study reports that the gene cluster is absent in *M. ruber* (Chen et al., 2008a). When the deduced amino acid sequence of the citrinin polyketide synthase (PKS) gene (GenBank accession no. BAD44749) was used as a search probe, a single relevant hit was found with the protein ID 469926 (gm1.3187_g), which shared 51.9% identity. This deduced protein shares 100% identity with the previously reported MAzP PKS from *M. pilosus* (Balakrishnan et al., 2013), and 97% identity with the MAzP PKS from *M. ruber* (GenBank accession no. JF832916). This gene is located at nt (nucleotide no.) 12,993-21,136 of the sequence scaffold 31, and the previously characterized *M. ruber* MAzP biosynthetic gene *MpigE* (Liu et al., 2014) can be found nearby at nt 28,591-29,845 of the scaffold 31. Inspection of the genetic organization in this region clearly indicates that this region corresponds to the MAzP biosynthetic

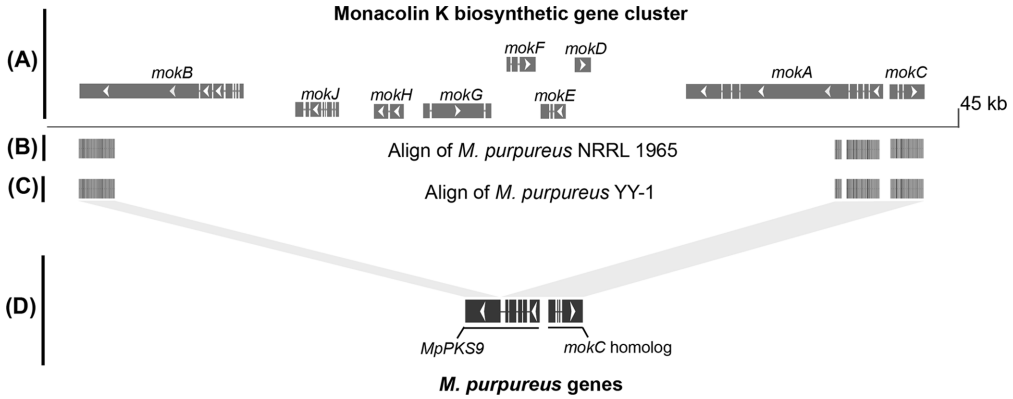

Fig. 1 Alignment of *M. purpureus* genomic regions with the MON-K (*mok*) gene cluster from *M. pilosus*. The genome sequence database in JGI was screened with the 45 kb *mok* gene cluster sequence (GenBank accession no. DQ176595), and the alignment regions were extracted for NCBI pairwise BLAST-N analyses. (A) Genetic organization of the *mok* gene cluster. The graphic representation was downloaded from NCBI, to be used after minor curation. (B and C) Alignments of *M. purpureus* NRRL 1965 (B) and YY-1 (C) genomic regions homologous to the *mok* cluster sequence. The default setting (expect threshold, 10; window size, 11; match/mismatch score, 2/-3) of Dialign BLAST-N was used for analysis and the graphic representation was downloaded for presentation. Aligned sequences shorter than 100-bp are omitted for clarity. (D) Gene organization of the *M. purpureus* regions used for alignment. We adopted the exon-intron architectures provided on the JGI webpage for drawing these two *M. purpureus* genes.

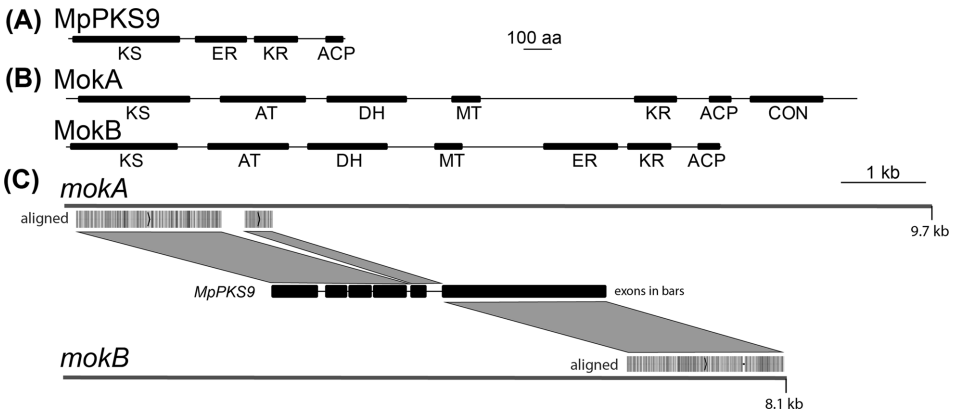

Fig. 2 Domain organizations of the deduced proteins of *MpPKS9* (A) and *mokA*/*mokB* (B). The domain architectures are deduced using NCBI's conserved domain analysis. Abbreviations of the domains are b-ketoacyl synthase (KS), acyltransferase (AT), dehydratase (DH) *C*-methyltransferase, enoylreductase (ER), ketoreductase (KR), acyl-carrier protein (ACP), and condensation (CON). (C) Homology alignment of *MpPKS9* with both *mokA* and *mokB*. The default setting (expect threshold, 10; window size, 11; match/mismatch score, 2/-3) of Dialign BLAST-N was used for analysis, and the graphic representation was downloaded for presentation.

gene cluster, confirming the absence of a citrinin gene cluster in *M. ruber* NRRL 1597.

Continuing our research on *Monascus* polyketide metabolic potential, we analyzed the genome sequences of *M. ruber* and *M. purpureus* for polyketide biosynthetic genes. The MON-K gene (*mok*) sequence of *M. pilosus* (GenBank accession no. DQ176595) was used for the genome search of highly-reducing PKS genes. The *mok* gene custer was composed of two highly-reducing PKS genes, *mokA* and *mokB*; the gene cluster ends with *mokB* and *mokC*, a cytochrome P450 gene (*lovA* for lovastatin) (Fig. 1A). The *M. ruber* genome harbors the MON-K gene cluster at scaffold 173 and 191. Scaffold 191 and 173 contain *mokB* and other *mok* genes, respectively. The *M. ruber* nucleotide sequence shares no less than 99% identity with the *mok* cluster sequence of *M. pilosus*. In contrary to *M. ruber*, the *mok* locus could not be located in the *M. purpureus* genome. Instead, only a homolog of *mokC* was identified in the *M. purpureus* genome, with 83% identity at nt 61,490-63,215 of scaffold 1 (Fig. 1B). When the *mokC*-nearby region of the *M. purpureus* genome sequence (scaffold 1; nt 57,211-64,505) was compared to the 45 kb *mok* gene cluster, an alignment was identified at both ends of the cluster but not in between (Fig. 1B). A BLAST-N (Altschul et al., 1997) generally found alignment with 70% or more identity, and the region other than the *mokC* region shared 75% identity overall. We deduced that there is no less than a 35 kb deletion in an ancestral *mok* gene cluster in *M. purpureus* NRRL 1596, and that *M. purpureus* NRRL 1596 is incapable of producing MON-K. However, we could not have excluded the possibility that this sequence deletion is a sequencing error. We thus examined the presence of MON-K gene cluster in the genome sequence of the *M. purpureus* industrial strain YY-1 (Yang et al., 2015). We downloaded the contigs sequences of the *M. purpureus* YY-1

genome (http://spxy.tust.edu.cn/duxj/index.html) and performed an off-line BLAST-N search to identify contig00763 (7,294 bp). This *M. purpureus* YY-1 contig sequence shared no less than 99% identity with the cognate sequence of *M. purpureus* NRRL 1596, and was aligned at both ends of *mok* cluster, similar to the *M. purpureus* NRRL 1596 sequence (Fig, 1C). This observation indicates that two *M. purpureus* strains lack a MON-K biosynthesis locus.

The aligned *M. purpureus* sequences, other than the *mokC* homolog, were found to encode a highly-reducing PKS gene (coined *MpPKS9* in our genome mining research), which was not included in the PKS genes deduced from the *M. purpureus* YY-1 genome sequence (Yang et al., 2015). A conserved domain analysis (Marchler-Bauer et al., 2015) implied that *MpPKS9* is non-functional due to a deletion inside (Fig. 2A). This gene apparently lacks in the regions for the acyltransferase, dehydratase, and methyltransferase domains (Fig. 2A and B). BLAST-N analysis indicated that the 5' and 3' regions of *MpPKS9* displayed homology to the 5' region of *mokA* and 3' region of *mokB*, respectively (Fig. 2C). We note that there is no alignment of *mokA* and *mokB* in this condition (data not shown). Thus, it is tempting to suggest that a deletion has occurred between *mokA* and *mokB* in an ancestral *mok* gene cluster of *M. purpureus*. Finally, we determined the nucleotide sequence of the *MpPKS9* region in our lab strain of *M. purpureus* Korean Agricultural Culture Collection (KACC) 42430. PCR primers designed with the *M. purpureus* NRRL 1596 sequence yielded DNA fragments of the expected sizes. Sanger sequencing of the resulting clones in pMD20 (Takara Bio Inc., Shiga, Japan) vector confirmed that the *MpPKS9* regions are identical in *M. purpureus* KACC 42430 and NRRL 1596 (data not shown).

Overall, we found that a type strain (NRRL 1596), an industrial strain (YY-1), and a laboratory strain (KACC 42430) of *M. purpureus* do not contain an intact *mok* gene cluster but instead harbor a remnant of the gene cluster. We therefore conclude that these strains are incapable of producing MON-K. A relatively low homology (75 to 80%) in this region between *M. purpureus* and *M. pilosus* suggests that *M. purpureus* is evolutionarily distant from both *M. ruber* and *M. pilosus*, as previously proposed, based on the distribution patterns of the citrinin biosynthetic genes (Chen et al., 2008a). Because some *M. purpureus* strains were reported as MON-K producers (Li et al., 2004), there might be two types of *M. purpureus* strains, one that supports MON-K biosynthesis and one that does not. We clarify that MON-K production is a strain-specific trait among *Monascus* species, providing a useful genetic marker for *M. purpureus* strains.

Acknowledgments This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2013R1A1A2059458).

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–402.

Balakrishnan B, Chen CC, Pan TM, and Kwon HJ (2014) Mpp7 controls regioselective Knoevenagel condensation during the biosynthesis of *Monascus* azaphilone pigments. *Tetrahedron Lett* **55**, 1640–3.

Balakrishnan B, Karki S, Chiu SH, Kim HJ, Suh JW, Nam B et al. (2013) Genetic localization and *in vivo* characterization of a Monascus azaphilone pigment biosynthetic gene cluster. *Appl Microbiol Biotechnol* **97**, 6337–45.

Chen YP, Tseng CP, Chien IL, Wang WY, Liaw LL, and Yuan GF (2008a) Exploring the distribution of citrinin biosynthesis related genes among *Monascus* species. *J Agric Food Chem* **56**, 11767–72.

Chen YP, Tseng CP, Liaw LL, Wang CL, Chen IC, Wu WJ et al. (2008b) Cloning and characterization of monacolin K biosynthetic gene cluster from *Monascus pilosus*. *J Agric Food Chem* **56**, 5639–46.

Feng Y, Shao Y, and Chen F (2012) *Monascus* pigments. *Appl Micorbiol Biotechnol* **96**, 1421–40.

Kennedy J, Auclair K, Kendrew SG, Park C, Vederas JC, and Hutchinson CR (1999) Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science* **284**, 1368–72.

Lee CL and Pan TM (2012) Development of *Monascus* fermentation technology for high hypolipidemic effect. *Appl Microbiol Biotechnol* **94**, 1449–59.

Li YG, Zhang F, Wang ZT, and Hu ZB (2004) Identification and chemical profiling of monacolins in red yeast rice using high-performance liquid chromatography with photodiode array detector and mass spectrometry. *J Pharm Biomed Anal* **35**, 1101–12.

Liu Q, Xie N, He Y, Wang L, Shao Y, Zhao H et al. (2014) *MpigE*, a gene involved in pigment biosynthesis in *Monascus ruber* M7. *Appl Microbiol Biotechnol* **9**8, 285–96.

Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY et al. (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Res* **43**, D222–6.

Patakova P (2013) *Monascus* secondary metabolites: production and biological activity. *J Ind Microbiol Biotechnol* **40**, 169–81.

Sakai K, Kinoshita H, Shimizu T, and Nihira T (2008) Construction of a citrinin gene cluster expression system in heterologous *Aspergillus oryzae*. *J Biosci Bioeng* **106**, 466–72.

Shimizu T, Kinoshita H, and Nihira T (2007) Identification and *in vivo* functional analysis by gene disruption of *ctnA*, an activator gene involved in citrinin biosynthesis in *Monascus purpureus*. *Appl Environ Microbiol* **73**, 5097–103.

Wang TH and Lin TF (2007) *Monascus* rice products. *Adv Food Nutr Res* **53**, 123–59.

Yang Y, Liu B, Du X, Li P, Liang B, Cheng X et al. (2015) Complete genome sequence and transcriptomics analyses reveal pigment biosynthesis and regulatory mechanisms in an industrial strain, *Monascus purpureus* YY-1. *Sci Rep* **5**:8331.

Yasuda M, Tachibana S, and Kuba-Miyara M (2012) Biochemical aspects of red koji and tofuyo prepared using *Monascus* fungi. *Appl Microbiol Biotechnol* **96**, 49–60.