Fungal Genetics and Biology 86 (2016) 58–70

Contents lists available at ScienceDirect

Fungal Genetics and Biology

journal homepage: www.elsevier.com/locate/yfgbi

Class of cyclic ribosomal peptide synthetic genes in filamentous fungi

Nozomi Nagano [a,b], Myco Umemura [c,*], Miho Izumikawa [d], Jin Kawano [c], Tomoko Ishii [e], Moto Kikuchi [e], Kentaro Tomii [a,b], Toshitaka Kumagai [f], Akira Yoshimi [g], Masayuki Machida [c], Keietsu Abe [g], Kazuo Shin-ya [a], Kiyoshi Asai [a,b,h]

[a] Biotechnology Research Institute for Drug Discovery, National Institute of Advanced Industrial Science and Technology (AIST), 2-4-7 Aomi, Koto-ku, Tokyo 135-0064, Japan
[b] Computational Biology Research Consortium (CBRC), 2-4-7 Aomi, Koto-ku, Tokyo 135-0064, Japan
[c] Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 17-2-1 Higashi-Nijo, Tsukisamu, Toyohira-ku, Sapporo, Hokkaido 062-8517, Japan
[d] Japan Biological Informatics Consortium, 2-4-7 Aomi, Koto-ku, Tokyo 135-0064, Japan
[e] Technology Research Association of Highly Efficient Gene Design, 17-2-1 Higashi-Nijo, Tsukisamu, Toyohira-ku, Sapporo, Hokkaido 062-8517, Japan
[f] Fermlab, Inc., 4-3-1-913 Shirakawa, Koto-ku, Tokyo 135-0021, Japan
[g] New Industry Creation Hatchery Center, Tohoku University, 6-6-10 Aoba, Aramaki, Aoba-ku, Sendai 980-8579, Japan
[h] Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8561, Japan

ARTICLE INFO

Article history:
Received 4 June 2015
Revised 10 November 2015
Accepted 14 December 2015
Available online 15 December 2015

Keywords:
Ribosomal peptide synthetic pathways
Aspergilli
Secondary metabolites
Genome analyses

ABSTRACT

Ustiloxins were found recently to be the first example of cyclic peptidyl secondary metabolites that are ribosomally synthesized in filamentous fungi. In this work, two function-unknown genes (*ustYa*/*ustYb*) in the gene cluster for ustiloxins from *Aspergillus flavus* were found experimentally to be involved in cyclization of the peptide. Their homologous genes are observed mainly in filamentous fungi and mushrooms. They have two "HXXHC" motifs that might form active sites. Computational genome analyses showed that these genes are frequently located near candidate genes for ribosomal peptide precursors, which have signal peptides at the N-termini and repeated sequences with core peptides for the cyclic portions, in the genomes of filamentous fungi, particularly *Aspergilli*, as observed in the ustiloxin gene cluster. Based on the combination of the *ustYa*/*ustYb* homologous genes and the nearby ribosomal peptide precursor candidate genes, 94 ribosomal peptide precursor candidates that were identified computationally from *Aspergilli* genome sequences were classified into more than 40 types including a wide variety of core peptide sequences. A set of the predicted ribosomal peptide biosynthetic genes was experimentally verified to synthesize a new cyclic peptide compound, designated as asperipin-2a, which comprises the amino acid sequence in the corresponding precursor gene, distinct from the ustiloxin precursors.

1. Introduction

Filamentous fungi produce secondary metabolites of various kinds such as polyketides, non-ribosomal peptides, alkaloids, and terpenes (Keller et al., 2005). Computational genome analyses using bioinformatics tools such as SMURF (Khaldi et al., 2010) and antiSMASH (Medema et al., 2011) can facilitate the identification of biosynthetic pathways for secondary metabolites including polyketides, non-ribosomal peptides and terpenes, based on their sequence motifs (Cacho et al., 2015; Keller et al., 2005; Wiemann and Keller, 2014). Recently, ribosomal peptides of various kinds, of which the core structures are encoded directly in precursor proteins and are ribosomally synthesized, have been reported as a new category of secondary metabolites from all three domains of life: archaea, bacteria, and eukaryotes (Arnison et al., 2013). For filamentous fungi, however, the peptide-type of secondary metabolites had been regarded as produced exclusively by non-ribosomal peptide synthetase (Keller et al., 2005). That inference of exclusive production prevailed until our reports described the biosynthetic gene clusters for ribosomal peptides, ustiloxins, from *Aspergillus flavus* (Umemura et al., 2014) and *Ustilaginoidea virens* (Tsukui et al., 2015).

Ustiloxins are toxic cyclic peptides that were originally identified from a pathogenic fungus for rice plant *U. virens* (Koiso et al., 1992, 1994, 1998). These peptides comprise Tyr(Y)-Val(V)/Ala(A)-

Abbreviations: EIC, extracted ion chromatogram; LC–MS, liquid chromatography–mass spectrometry; RiPS, ribosomal peptide synthetic; ust-RiPS, ustiloxin-type ribosomal peptide synthetic.

* Corresponding author.
E-mail address: umemura-m@aist.go.jp (M. Umemura).

http://dx.doi.org/10.1016/j.fgb.2015.12.010
1087-1845/

Ile(I)-Gly(G), a tetrapeptide that is cyclized at the sidechains of Tyr and Ile. The precursor protein for ustiloxin B from *A. flavus* includes 16 YAIG motifs (Umemura et al., 2014), whereas the counterpart protein for ustiloxins has five YVIG motifs and three YAIG motifs, respectively, for ustiloxins A and B from *U. virens* (Tsukui et al., 2015). These characteristics by which the core motif of peptide compounds is repeated in the precursor protein are observed in no other ribosomal peptide in any other organism. The ustiloxin gene clusters from *A. flavus* and *U. virens* respectively contain 15 and 13 genes, which encode the precursor protein, a transcription factor, a transporter, and several proteins that are presumably enzymes (Umemura et al., 2014; Tsukui et al., 2015). Although most of these hypothetical enzymes are homologous to the function-elucidated enzymes, two remaining genes, *ustYa* and *ustYb*, which are mutually homologous, are homologous to no function-known protein (Umemura et al., 2014).

This work demonstrates that the homologues of these two function-unknown genes (*ustYa* and *ustYb*), observed specifically in filamentous fungi (*Ascomycota–Pezizomycotina*) and mushrooms (*Basidiomycota–Agaricomycotina*), can be involved in the synthesis of ribosomal peptides in filamentous fungi. These *ustYa/Yb* homologous genes are frequently located near candidate genes for the ribosomal peptide precursor proteins that contain signal peptides and repeated motifs of peptides (ustiloxin-type ribosomal peptide synthetic (ust-RiPS) precursor gene), as observed in the ustiloxin precursor genes from *A. flavus* and *U. virens*. Based on the combination of the *ustYa/Yb* homologous genes and ust-RiPS precursor gene candidates in *Aspergilli* genome sequences, 94 sets of precursor candidates and *ustYa/Yb* homologues were computationally identified and classified. Finally, one of these identified sets of the precursor candidate genes and *ustYa/Yb* homologues in *A. flavus* was verified experimentally to be involved in the synthesis of a novel cyclic peptide compound, asperipin-2a, which was designated as a type-2a ribosomal peptide from *Aspergilli*. Asperipin-2a comprises the amino acid sequence in the repeated core peptide in the corresponding precursor protein, which is distinct from that for ustiloxins. The results strongly suggest that ust-RiPS precursor genes accompanied with *ustYa/Yb* homologues are involved in secondary metabolic pathways in fungi, in addition to those for non-ribosomal peptides, polyketides, alkaloids, and terpenes.

2. Material and methods

2.1. Strains and sequence

Aspergillus flavus strain CA14 Δ*ku70* Δ*pyrG* Δ*niaD* was used throughout this study with the gene-annotated genome sequence of *A. flavus*, GenBank EQ963472–EQ966232 (Yu et al., 2008), to design primers and to analyze the sequences. Deletion mutants for the 13 ustiloxin biosynthetic genes confirmed by RNA-seq analysis (NCBI accession no. BR001206; Tsukui et al., 2015), *ustC* (AFLA_094960), *ustD* (AFLA_095040), *ustO* (AFLA_094940), *ustF1* (AFLA_094950), *ustA* (AFLA_094980), *ustYa* (AFLA_094990), *ustH* (AFLA_095030), *ustF2* (AFLA_095050), *ustQ* (AFLA_095060), *ustT* (AFLA_095070), *ustR* (AFLA_095090), *ustM* (AFLA_095100), and *ustS* (AFLA_095110), and the *pyrG* revertant as control, which were previously prepared (Umemura et al., 2013, 2014), were used. Although the *ustR* gene comprises AFLA_095080 and AFLA_095090, the deletion mutant of AFLA_095090 was regarded as the deletion mutant of *ustR* because it mainly includes the domain of a fungal C6-type transcription factor (Umemura et al., 2014). The *pyrG* marker was reverted in the control strain to keep auxotrophy the same among the deletion mutants and the control strain (Umemura et al., 2013, 2014).

2.2. Gene disruption and transformation

Gene coding regions of *ustP* and *ustYb* were corrected to a great degree using RNA-seq analysis (Tsukui et al., 2015). Thereby the deletion mutants for the corrected *ustP* and *ustYb* genes were prepared by fusion PCR and protoplast transformation with *pyrG* as the selectable marker and the modified Cre-loxP marker recycling system (Mizutani et al., 2012), fundamentally as described in a previous report (Umemura et al., 2014). As a candidate of a newly identified RiPS cluster in *A. flavus*, AFLA_041390 and AFLA_041400 as an *ustYa/Yb* homologue and a precursor gene, respectively, were disrupted as well. The used primer pairs are presented in Table S1. DNA fragments used for cloning and PCR were amplified using KOD – Plus – DNA polymerase (Toyobo Co. Ltd., Japan).

2.3. Sequence analyses of UstYa and UstYb; Distribution in organisms and possible active-site residues

To detect homologous protein sequences for UstYa and UstYb, PSI-BLAST search (e-value 1e−4; threshold for inclusion in multipass model 0.001; Number of iteration 15) (Altschul et al., 1997) was conducted against the UniProtKB database (ver. 2013_08) (Magrane and UniProt Consortium, 2011), which includes protein sequence data from archaea, bacteria, and eukaryotes in addition to those from viruses. The species of the detected homologous proteins were checked for the corresponding UniProtKB entries. Multiple sequence alignment was performed for the top three sequences each for UstYa and UstYb using the Clustal Omega program (Sievers et al., 2011). The default option was used for the Clustal Omega alignment of close homologues. Signal peptide detection was conducted using SignalP 4.1 software (Nielsen et al., 1997; Petersen et al., 2011). Transmembrane helix detection was also done using TMHMM (Möller et al., 2001). To detect the disordered regions, DISOPRED2 (Ward et al., 2004) was used for sequences of UstYa and UstYb.

2.4. Detection of ustiloxin-type ribosomal peptide biosynthetic gene candidates in fungal genomes

To identify ust-RiPS gene cluster candidates in filamentous fungi, sequence analyses of AspGD, which is a database for genomes of *Aspergilli* species (Cerqueira et al., 2014), were conducted. A flow chart showing the analyses is presented in Fig. S1A. The genome of *Fusarium verticillioides* from GenBank CM000578–CM000588 (Ma et al., 2010) was also analyzed similarly.

2.4.1. Detection of ustYa/Yb homologous gene products in Aspergilli genomes by PSI-BLAST

PSI-BLAST search (e-value 1e−3; threshold for inclusion in multipass model 0.01) (Altschul et al., 1997) was conducted using UstYa and UstYb amino acid sequences against the amino acid sequence dataset for the AspGD database (ver. February 14, 2014) (Cerqueira et al., 2014). They were defined as *ustYa/Yb* homologues if the sequences are hit by either UstYa or UstYb. The sequence search procedure was iterated 22 times collecting 257 homologous sequences for UstYa, whereas it was iterated only five times collecting 208 homologues for UstYb until the convergence. To exclude possible false positives, the regions of the hit sequences were also considered so that the hit sequences contain the plausible active-site region, particularly two HXXHC motifs described below, of the query sequences (residues 165–196 for UstYa and 147–201 for UstYb). Consequently, in the final stage, out of 257 sequences hit by either UstYa or UstYb, 13 hit sequences that do not contain the active-site regions were discarded. Of the 244 detected sequences, e-values for 183 sequences in the last iteration of PSI-BLAST search were <1e−10, although the highest value