RESEARCH ARTICLE

Functional Operons in Secondary Metabolic Gene Clusters in *Glarea lozoyensis* (Fungi, Ascomycota, Leotiomycetes)

Qun Yue,[a,b] Li Chen,[a,b,c] Yan Li,[b] Gerald F. Bills,[b] Xinyu Zhang,[a] Meichun Xiang,[a] Shaojie Li,[a] Yongsheng Che,[d] Chengshu Wang,[e] Xuemei Niu,[b,f] Zhiqiang An,[b] Xingzhong Liu[a]

State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China[a]; Texas Therapeutics Institute, the Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center at Houston, Houston, Texas, USA[b]; University of Chinese Academy of Sciences, Beijing, China[c]; State Key Laboratory of Toxicology & Medical Countermeasures, Beijing Institute of Pharmacology & Toxicology, Beijing, China[d]; Key Laboratory of Insect Developmental and Evolutionary Biology, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China[e]; Laboratory for Conservation and Utilization of Bio-Resources and Key Laboratory for Microbial Resources of the Ministry of Education, Yunnan University, Kunming, Yunnan, China[f]

Q.Y., L.C., and Y.L. contributed equally to this work.

ABSTRACT **Operons are multigene transcriptional units which occur mostly in prokaryotes but rarely in eukaryotes. Protein-coding operons have not been reported in the Fungi even though they represent a very diverse kingdom of organisms. Here, we report a functional operon involved in the secondary metabolism of the fungus *Glarea lozoyensis* belonging to Leotiomycetes (Ascomycota). Two contiguous genes, *glpks3* and *glnrps7*, encoding polyketide synthase and nonribosomal peptide synthetase, respectively, are cotranscribed into one dicistronic mRNA under the control of the same promoter, and the mRNA is then translated into two individual proteins, GLPKS3 and GLNRPS7. Heterologous expression in *Aspergillus nidulans* shows that the GLPKS3-GLNRPS7 enzyme complex catalyzes the biosynthesis of a novel pyrrolidinedione-containing compound, xenolozoyenone (compound 1), which indicates the operon is functional. Although it is structurally similar to prokaryotic operons, the *glpks3-glnrps7* operon locus has a monophylogenic origin from fungi rather than having been horizontally transferred from prokaryotes. Moreover, two additional operons, *glpks28-glnrps8* and *glpks29-glnrps9*, were verified at the transcriptional level in the same fungus. This is the first report of protein-coding operons in a member of the Fungi.**

IMPORTANCE **Operons are multigene transcriptional units which occur mostly in prokaryotes but rarely in eukaryotes. Three operon-like gene structures for secondary metabolism that were discovered in the filamentous fungus *Glarea lozoyensis* are the first examples of protein-coding operons identified in a member of the Fungi. Among them, the *glpks3-glnrps7* operon is responsible for the biosynthesis of xenolozoyenone, which is a novel tetramic acid-containing compound. Although structurally similar to prokaryotic operons, the *glpks3-glnrps7* operon locus did not result from horizontal gene transfer from prokaryotes. In addition, operonlike structures have been predicted *in silico* to be common in other fungi. The common occurrence and operonlike structure in fungi provide evolutionary insight and essential data for eukaryotic gene transcription.**

Received 24 April 2015 **Accepted** 14 May 2015 **Published** 16 June 2015

Citation Yue Q, Chen L, Li Y, Bills GF, Zhang X, Xiang M, Li S, Che Y, Wang C, Niu X, An Z, Liu X. 2015. Functional operons in secondary metabolic gene clusters in *Glarea lozoyensis* (Fungi, Ascomycota, Leotiomycetes). mBio 6(3):e00703-15. doi:10.1128/mBio.00703-15.

Editor B. Gillian Turgeon, Cornell University

Address correspondence to Zhiqiang An, zhiqiang.an@uth.tmc.edu, and Xingzhong Liu, liuxz@im.ac.cn.

Genes that cooperate in a particular biological process often require coordinated regulation of their expression (1). In prokaryotic genomes, many functionally coupled genes are grouped together within an operon that allows the genes to be cotranscribed under the control of a single regulatory element, which results in a single polycistronic mRNA. In contrast, eukaryotic genes are usually transcribed individually into monocistronic mRNA. Although operons were previously thought to exist solely in prokaryotes, clustered protein-coding genes with coexpression behave as operons in some eukaryotes (2), and rRNA genes are often grouped in operons. Except for the operons transcribed into rRNAs that do not encode proteins, the other eukaryotic operons reported so far are divided into two broad types. The first type of eukaryotic operon is transcribed to produce polycistronic transcripts that are subsequently converted into mature monocistronic mRNAs. This type of operon was first reported in the nematode *Caenorhabditis elegans* and subsequently in other nematodes and some early-diverging chordates (3–5). A global analysis of the *C. elegans* genome showed that about 15% of all *C. elegans* genes are in operons (6). The second type of eukaryotic operon is transcribed as dicistronic mRNAs that are translated as in prokaryotic operons. Examples of the latter type have been found in plants, flies, and mammals, encoding γ-glutamyl kinase and γ-glutamyl phosphate reductase, sugar receptors, and a candidate gene for a role in the imprinted Prader-Willi syndrome, respectively (7–9). However, operons with protein-coding genes have not been previously reported in the Fungi even though they represent a kingdom with more than 100,000 described species and potentially more than a million awaiting discovery.

The filamentous fungus *Glarea lozoyensis* (ATCC 20868) pro-

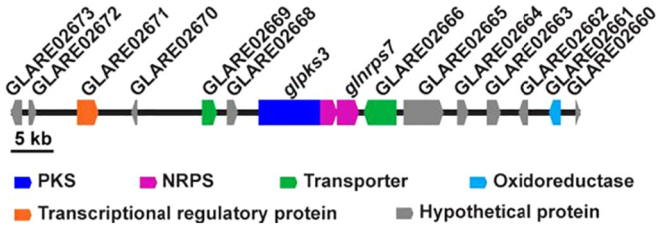

FIG 1 Characterization of the *glpks3-glnrps7* gene cluster in *Glarea lozoyensis*. The *glpks3-glnrps7* locus is located in a 66.6-kb gene cluster that includes two putative transporter genes, one putative oxidoreductase gene, and one putative transcriptional regulatory protein gene.

duces pneumocandin B_0, which is the starting molecule for the production of the antifungal drug caspofungin acetate (10). To identify the gene cluster, comprised of a nonribosomal peptide synthetase (NRPS) and a polyketide synthase (PKS), that is required for the biosynthesis of pneumocandin B_0, we sequenced the genome of *G. lozoyensis* and annotated the gene structures for secondary metabolite biosynthesis (10). Among six PKS and NRPS gene clusters, five are predicted to be involved in the biosynthesis of hybrid polyketide-nonribosomal peptide molecules containing 1 amino acid residue. cDNA analysis revealed that three of these five polyketide synthase-nonribosomal peptide synthetase (PKS-NRPS) hybrid genes are transcribed as operons. After analysis at the protein and metabolite levels, one of them, containing the genes *glpks3* and *glnrps7*, is confirmed to be transcribed as an operon, and the two resultant proteins catalyze the biosynthesis of a novel secondary metabolite.

RESULTS

The *glpks3* and *glnrps7* genes are transcribed as an operon. The *glpks3-glnrps7* locus, annotated as a PKS-NRPS hybrid gene, is part of a typical secondary metabolic gene cluster (Fig. 1). To investigate the transcription of *glpks3-glnrps7*, we used 5′- rapid amplification of cDNA ends (5′-RACE), 3′-RACE, and reverse transcription (RT)-PCR to amplify seven overlapping cDNA fragments to obtain the whole mRNA transcript. Each of the seven overlapping cDNA fragments spans at least one intron in the corresponding genomic DNA sequence to ensure that they were amplified from cDNA templates. By comparing the genomic DNA and the cDNA sequences, we confirmed that *glpks3-glnrps7* is transcribed as a single 12,562-nucleotide (nt) transcript (Fig. 2A and B). We assessed the cDNA sequence for potential open reading frames (ORFs) and confirmed two distinct ORFs separated by a 26-bp intergenic sequence (Fig. 2C and D). *glpks3* contains a 9,021-bp ORF with six introns, and *glnrps7* contains a 2,664-bp intronless ORF. The 26-bp intergenic sequence between *glpks3* and *glnrps7* lacks a polyadenylation signal, which is required for 3′-end cleavage and polyadenylation of pre-messenger RNA in eukaryotic protein-coding genes (11). In addition, neither a transcriptional termination site for *glpks3* nor a transcription initiation site for *glnrps7* was detected. Taken together, these results confirm that *glpks3* and *glnrps7* are cotranscribed into one dicistronic mRNA.

Since the *glpks3-glnrps7* locus contains two ORFs, two distinct proteins are expected to be translated from the single 12,562-nt transcript. We carried out two experiments to confirm the translation of two proteins from the 12,562-nt transcript. First, a

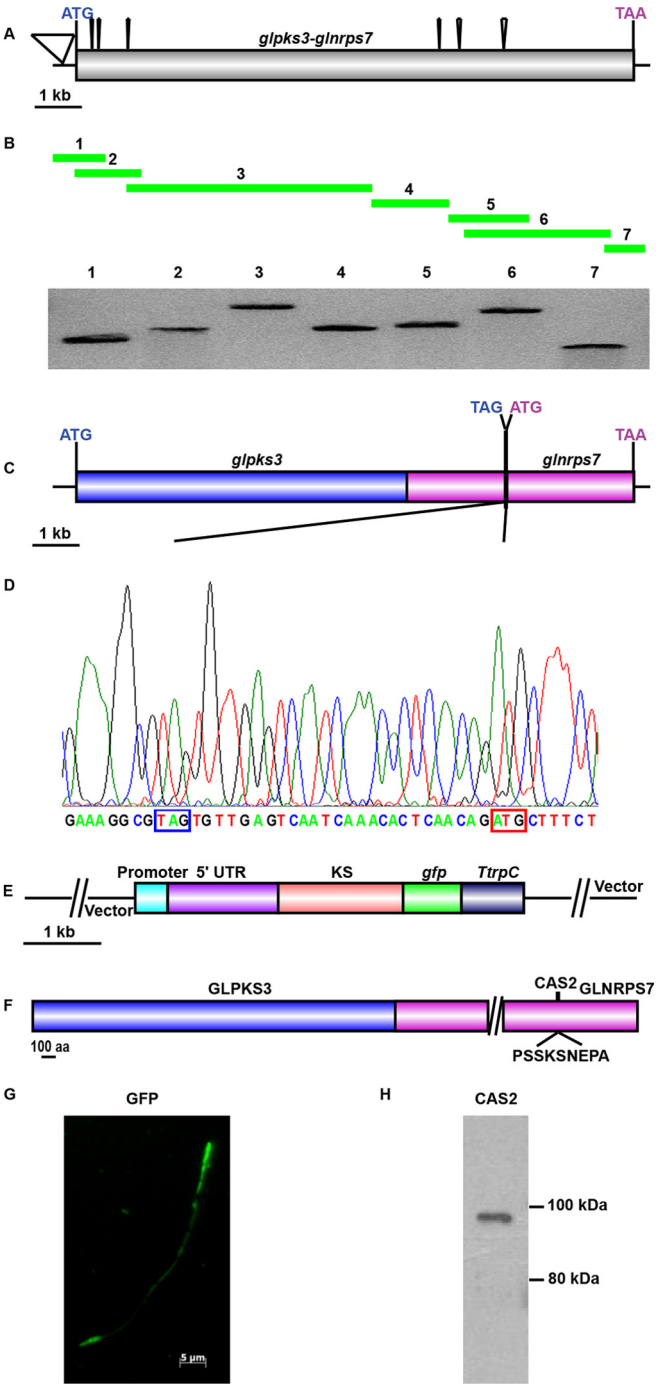

FIG 2 *glpks3-glnrps7* operon gene organization in *Glarea lozoyensis*. (A) Characterization of the *glpks3-glnrps7* locus. Vertical pointers indicate intron positions. (B) Transcription of *glpks3-glnrps7* in *G. lozoyensis*. Fragments 1 and 7 were obtained by RACE-PCR, and the other fragments were obtained by RT-PCR. The whole transcript was assembled based on the 7 overlapping cDNA fragments. (C) Characterization of the whole *glpks3-glnrps7* transcript. (D) Sequencing chromatograph of the 26-bp intergenic gap between *glpks3* and *glnrps7*. The blue rectangle indicates the first stop codon, and the red rectangle indicates the second start codon. (E) GFP expression vector pAg1-H3L-KS2-GFP-TtrpC. UTR, untranslated region; KS, ketosynthase domain. (F) Predicted GLPKS3-GLNRPS7 proteins. Polyclonal antibody CAS2, prepared by immunizing a rabbit with the peptide antigen PSSKSNEPA, was used to detect the expression of GLNRPS7. (G) GFP expression in the mycelium in the KS::GFP *G. lozoyensis* strain as observed by fluorescence microscopy. (H) Western blotting detection of GLNRPS7.

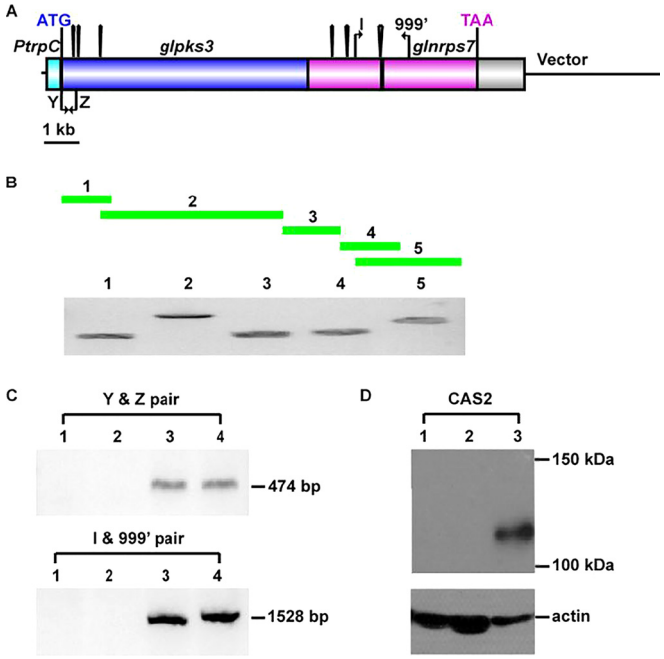

FIG 3 Verifying the *glpks3-glnrps7* operon gene organization in *Aspergillus nidulans*. (A) Heterologous expression vector pAg1-PtrpC-glpks3-glnrps7. Primer Y and Z and I and 999′ were used to detect transcription of *glpks3-glnrps7* in *A. nidulans* by RT-PCR. The RT-PCR products were 474 bp and 1,528 bp, respectively. (B) Transcription of *glpks3-glnrps7* in *PtrpC-glpks3-glnrps7* transformant of *A. nidulans*. All fragments were obtained by RT-PCR. (C) RT-PCR analysis of *glpks3-glnrps7*. Lane 1, negative control *A. nidulans* A773; lane 2, negative control *glpks3-glnrps7* transformant of *A. nidulans*; lane 3, *PtrpC-glpks3-glnrps7* transformant of *A. nidulans*; lane 4, positive control *Glarea lozoyensis*. Primer sets Y/Z and I/999′ yielded a 474-bp RT-PCR product and a 1,528-bp RT-PCR product, respectively, in the *PtrpC-glpks3-glnrps7* transformant of *A. nidulans*. (D) Western blotting analysis of GLNRPS7. Lane 1, negative control *A. nidulans* A773; lane 2, negative control *glpks3-glnrps7* transformant of *A. nidulans*; lane 3, *PtrpC-glpks3-glnrps7* transformant of *A. nidulans*.

genomic DNA fragment spanning the promoter sequence and the region annotated to encode the ketosynthase domain (KS) of the *glpks3* gene was fused downstream with a green fluorescent protein (GFP) reporter gene (Fig. 2E). The reporter plasmid was introduced into the *G. lozoyensis* wild-type strain by *Agrobacterium tumefaciens*-mediated transformation. A GFP signal was detected in the transformants (Fig. 2G), indicating that the transcript was translated. Second, we attempted to detect the predicted GLNRPS7 protein, containing 887 amino acids and with a molecular mass of about 98 kDa, by Western blotting. A rabbit polyclonal antibody (CAS2) was generated against a peptide containing residues 361 to 369 of GLNRPS7 (Fig. 2F). A protein band with a molecular mass close to that of the predicted GLNRPS7 protein of 98 kDa was detected in the mycelial homogenate of the *G. lozoyensis* wild-type strain by Western blotting using the CAS2 antibody (Fig. 2H). These results demonstrated that the single *glpks3-glnrps7* transcript is translated into two distinct proteins.

To confirm that transcription of the *glpks3-glnrps7* operon is controlled by a single promoter, the *glpks3-glnrps7* locus without its endogenous promoter was cloned and introduced by protoplast transformation into the genetic model fungus *Aspergillus nidulans* under the control of the *A. nidulans trpC* promoter (*PtrpC*) (Fig. 3A). We mapped the *PtrpC-glpks3-glnrps7* transcript in the *A. nidulans* transformant by amplifying five overlapping cDNA fragments, and the results showed that the *PtrpC-glpks3-glnrps7* gene construct was transcribed into a single transcript which was identical to the transcript of *glpks3-glnrps7* in *G. lozoyensis* (Fig. 3B). The transcription of *glpks3* and *glnrps7* in the heterologous host *A. nidulans* was further confirmed by RT-PCR using two primer sets (Fig. 3A). One pair (5′-Y/3′-Z) spanned the first intron of *glpks3*, and the other pair (5′-I/3′-999′) spanned the last intron of *glpks3* and the intergenic region plus part of *glnrps7*. The results showed that the *glpks3* and *glnrps7* genes were transcribed in the *A. nidulans* strain transformed with *PtrpC-glpks3-glnrps7* and the *G. lozoyensis* wild-type strain but not in the *A. nidulans* parent strain or the *A. nidulans* strain transformed with the promoterless *glpks3-glnrps7* construct (Fig. 3C). These results confirmed that transcription of the *glpks3* and *glnrps7* operon is controlled by a single promoter. We also observed a protein of approximately 120 kDa only in the *PtrpC-glpks3-glnrps7 A. nidulans* transformant by Western blotting using antibody CAS2 (Fig. 3D). The larger molecular mass of the GLNRPS7 protein in *A. nidulans* may be the result of heavier glycosylation in the heterologous host. Based on these results, we conclude that the *glpks3* and *glnrps7* genes are cotranscribed into one dicistronic mRNA under the control of a single promoter and the resultant transcript is then translated into two individual proteins, GLPKS3 and GLNRPS7, proving that the *glpks3-glnrps7* locus belongs to the second type of eukaryotic operon in which the dicistronic mRNAs are translated as in prokaryotic operons.

Functional analysis of the *glpks3-glnrps7* gene cluster. The *glpks3-glnrps7* locus is clustered with two putative transporter genes, one putative oxidoreductase gene, and one putative transcriptional regulatory gene (Fig. 1), which constitutes a typical organization for a gene cluster for the biosynthesis of a secondary metabolite. Based on the catalytic domain structures of GLPKS3-GLNRPS7, it is predicted that the gene cluster is responsible for the biosynthesis of a polyketide chain linked to 1 amino acid residue. In order to detect this compound, we generated *G. lozoyensis* mutant strains in which the DNA fragment from 8,512 to 9,946 bp downstream from the translation start codon of *glpks3* (including the 26-bp gap and partial region annotated as an adenylation domain [A]) was deleted by *Agrobacterium tumefaciens*-mediated transformation to inactivate GLPKS3-GLNRPS7 (Δ*glpks3-glnrps7*) (see Fig. S1 in the supplemental material). By comparing the metabolite profiles of the wild-type and Δ*glpks3-glnrps7* strains, we hoped to identify the compound encoded by the *glpks3-glnrps7* gene cluster. Repeated attempts under different fermentation conditions did not reveal any differences between the high-performance liquid chromatography-mass spectrometry (HPLC-MS) profiles of the wild-type strain and the Δ*glpks3-glnrps7* deletion strain. As an alternative strategy, we compared the HPLC-MS profiles of the *A. nidulans* parent strain and the *A. nidulans* strain transformed with the *PtrpC-glpks3-glnrps7* gene construct. One new and unique peak was observed in the fermentation broth of the *PtrpC-glpks3-glnrps7* transformant of *A. nidulans* but not in the fermentation broth of the *A. nidulans* parent strain or the *A. nidulans* strain transformed with the *glpks3-glnrps7* construct without a promoter (Fig. 4A). We purified the target compound, referred to as compound 1, and elucidated its structure with the high-resolution electrospray ionization-mass spectrometry (HRESIMS) and one-dimensional and two-dimensional nuclear magnetic resonance (NMR) spectra (Fig. 4B; see also Fig. S2).

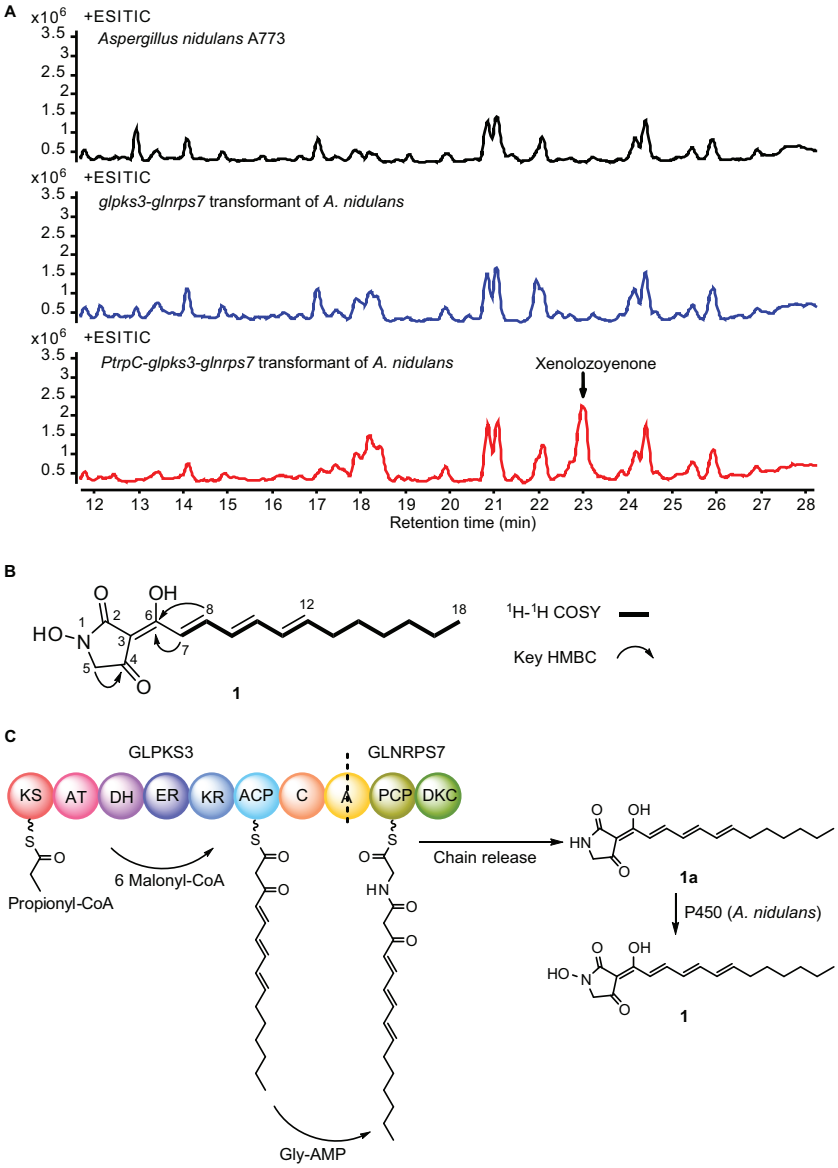

FIG 4 *Glarea lozoyensis* GLPKS3-GLNRPS7 biosynthetic product xenolozoyenone biosynthesized in the heterologous host *Aspergillus nidulans*. (A) HPLC-MS profiles for negative control *A. nidulans* A773, negative control *glpks3-glnrps7* transformant of *A. nidulans*, and *PtrpC-glpks3-glnrps7* transformant of *A. nidulans*. The peak corresponding to xenolozoyenone was only observed in the cultures of the *PtrpC-glpks3-glnrps7* transformant of *A. nidulans*. ESITIC, electrospray ionization total ion current. (B) Structure of xenolozoyenone (compound 1) and its ^{1}H–^{1}H COSY and key HMBC correlations. The ^{1}H–^{1}H COSY showed that compound 1 contains an isolated spin system of C-7–C-18, while the key HMBC correlations from H-8 and H-7 to C-6 indicated that the side chain was connected to the 2,4-pyrrolidinedione moiety at C-6. (C) Proposed route for xenolozoyenone biosynthesis. The PKS portion is responsible for the formation of the linear polyketide chain, and the NRPS portion for the extension of the polyketide chain with an activated glycine. The linear polyketide-nonribosomal peptide chain is then released, resulting in the formation of compound 1a. Finally, compound 1a could be modified by tailoring enzymes (e.g., cytochrome P450) in *A. nidulans* to yield compound 1. KS, ketosynthase domain; CoA, coenzyme A; AT, acyltransferase domain; DH, dehydratase domain; ER, enoylreductase domain; KR, β-ketoacylreductase domain; ACP, acyl carrier protein; C, condensation domain; A, adenylation domain; PCP, peptidyl carrier protein; DKC, Dieckmann cyclase domain.

It is a novel tetramic acid-containing metabolite and is given the trivial name xenolozoyenone. The prefix "xeno" refers to the biosynthetic gene's expression in a foreign host. Xenolozoyenone is a polyketide chain linked to an amino acid residue, which is consistent with the expected biosynthetic product of the *glpks3-glnrps7* gene cluster. Collectively, these results suggest that the *glpks3-glnrps7* operon is functional at the transcriptional, translational, and chemical levels.

Evolutionary origin of GLPKS3-GLNRPS7 operon. Operons are common in prokaryotes. In order to clarify the evolutionary origin of the *glpks3-glnrps7* operon, we constructed a phylogenetic tree based on the KS domain of GLPKS3, 30 fungal PKSs, and 9 bacterial PKSs (Fig. 5; see also Fig. S3 and S4 and Table S1 in the supplemental material). The 30 fungal PKSs and 9 bacterial PKSs have been linked to products with known chemical structures (Table S1) and, thus, were proven to be functional PKSs. GLPKS3 and

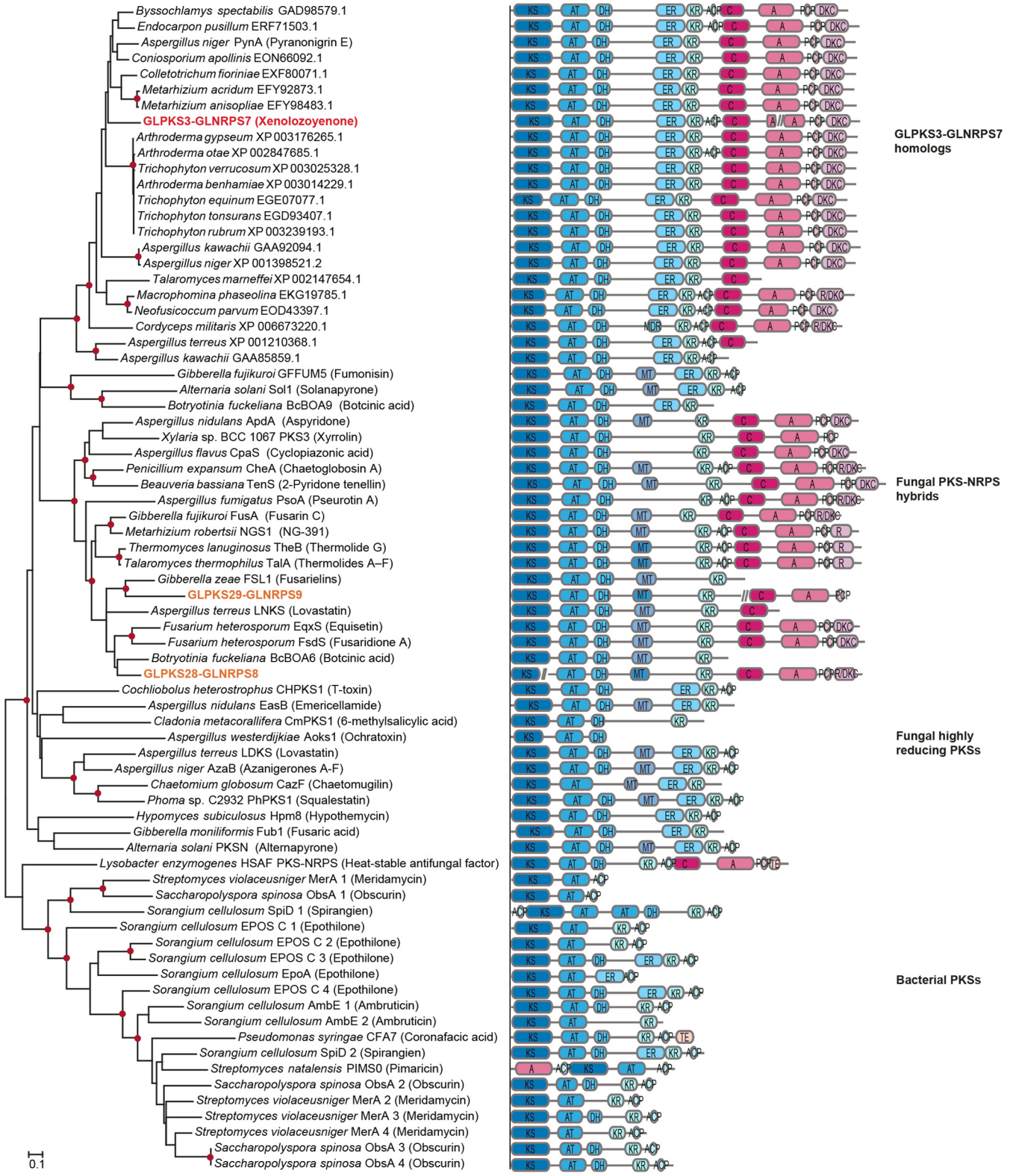

FIG 5 Domain prediction and phylogenetic analysis of ketosynthase domain (KS) of GLPKS3-GLNRPS7 and its orthologs, GLPKS28-GLNRPS8, GLPKS29-GLNRPS9, 30 characterized fungal polyketide synthases (PKSs), and 9 characterized bacterial PKSs. The domains were annotated by anti-SMASH or visually in multiple alignments. The genealogy of PKSs was inferred by maximum-likelihood analysis of the aligned predicted amino acid sequences of the KS domains. The branch length indicates the number of inferred amino acid changes. Red dots indicate branch nodes with >70% support. AT, acyltransferase domain; DH, dehydratase domain; MT, methyltransferase; MDR, medium-chain reductase/dehydrogenase domain; ER, enoylreductase domain; KR, β-ketoacylreductase domain; ACP, acyl carrier protein; C, condensation domain; A, adenylation domain; PCP, peptidyl carrier protein; R, reductive domain; DKC, Dieckmann cyclase domain; TE, thioesterase domain.

the 30 fungal PKSs were grouped together (bootstrap value of 92%), and the clade was distantly related to and distinct from the bacterial PKSs. Thus, the *glpks3-glnrps7* locus and its fungal orthologs appear to have a monophyletic origin among fungal PKSs, and the *glpks3-glnrps7* locus did not result from horizontal gene transfer from prokaryotes. Among the fungal PKSs, GLPKS3 and its fungal orthologs were grouped in a monophyletic clade with a high bootstrap value (100%). In this clade, 26 proteins were from 13 genera belonging to four classes of Pezizomycotina, which suggests that these kinds of enzymes are widespread and possibly serve a basic function in these fungi.

Two other *G. lozoyensis* gene loci transcribed as operons. In addition to *glpks3-glnrps7*, the *G. lozoyensis* genome revealed four more genes (*glpks26-nrps*, *glpks27-nrps*, *glpks28-nrps*, and *glpks29-nrps*) that are predicted to be responsible for the biosynthesis of polyketide linked to 1 amino acid residue (10). To find out if the occurrence of operons is common among these secondary metabolic genes in *G. lozoyensis*, we determined the cDNAs of the other four PKS-NRPS hybrid genes. Whole transcripts assembled with overlapping cDNA fragments obtained by RACE-PCR and RT-PCR revealed that *glpks28-nrps* and *glpks29-nrps* possess gene organization patterns similar to that of the *glpks3-glnrps7* locus (Fig. 6). Because not every cDNA fragment spans an intron, one primer set, PF5-1 and Ourer5 (located in the third intron of *glpks29-nrps*), was used as a genomic DNA contamination control. No PCR product was amplified by this primer set with the template cDNA, indicating there was no genomic DNA contamination in the RNA samples used for cDNA amplification. The *glpks28-nrps* locus was transcribed as a dicistronic transcript with an intergenic sequence of 443 nt, while the *glpks29-nrps* locus also was transcribed as a dicistronic transcript with an intergenic sequence of 179 nt. No polyadenylation signal sequences or promoter elements were detected in the intergenic region. Moreover, no transcriptional termination site for the upstream gene or transcription initiation site for the downstream gene were detected in the intergenic regions. These results demonstrated that the *glpks28-nrps* and *glpks29-nrps* loci are organized as operons. Based on these results, we renamed these two loci as *glpks28-glnrps8* and *glpks29-glnrps9*. *glpks28* contained a 966-bp ORF, and *glnrps8* contained a 10,822-bp ORF. *glpks29* had a 9,212-bp ORF, and *glnrps9* had a 3,439-bp ORF. Among the four genes, *glpks28* contains no introns, while *glnrps8*, *glpks29*, and *glnrps9* contain multiple introns. These two loci are organized as typical secondary metabolic gene clusters which are predicted to be responsible for the biosynthesis of compounds consisting of a polyketide chain linked to an amino acid residue.

DISCUSSION

rRNA operons are widespread in eukaryotes, including the Fungi. However, protein-coding operons have not been reported in the Fungi. A dicistronic mRNA of *CRG1* and *PUT1* was reported in the filamentous fungus *Cercospora nicotianae*, but these two genes are controlled by separate promoters (12). Therefore, this locus does not fulfill the definition of an operon (2). We demonstrated experimentally in this study that three of the five PKS-NRPS loci in *G. lozoyensis* are operons. To our knowledge, the three operons in *G. lozoyensis* are the first examples of eukaryotic protein-coding operons identified in a member of the fungal kingdom.

The majority of known eukaryotic operons belong to the first type of eukaryotic operons, in which the polycistronic transcripts are processed into mature monocistronic mRNAs (2). Although the mechanism for transcript maturation in *C. elegans* operons is different from that in the *Drosophila melanogaster CheB42a/llz* operon, the polyadenylation signal in the intergenic region is crucial for the 3′-end formation in all of these operons (6, 13). No polyadenylation signal was found in the intergenic regions of the *glpks3-glnrps7*, *glpks28-glnrps8*, and *glpks29-glnrps9* operons, suggesting that the operons in *G. lozoyensis*, like the *stonedA/stonedB* operon in *D. melanogaster* (14), belong to the second type of eukaryotic operons, in which the dicistronic mRNAs are translated as in prokaryotic operons.

It has been suggested that some of the second type of eukaryotic operons had originated from bacteria via horizontal gene transfer (8). However, phylogenetic analysis of the *glpks3-glnrps7*, *glpks28-glnrps8*, and *glpks29-glnrps9* operons showed that they do not resemble prokaryotic genes (Fig. 5). Not only do the three operons contain multiple introns, but they also have GC content (46 to 50%) consistent with that the resident genome (46%) (10). These results suggest that these operons were not products of horizontal gene transfer from bacteria. We also constructed phylogenetic trees for the KS and A domains of GLPKS3-GLNRPS7 and their orthologs (see Fig. S5 in the supplemental material). The proteins containing both KS and A orthologs formed distinct clades with GLPKS3-GLNRPS7 in the two trees. However, the A domain tree appears to be in conflict with the KS tree and established ascomycete phylogeny, while the KS tree does not significantly conflict with the established phylogeny. These results suggest that these proteins may have descended from a common ancestor, but the NRPS module and the PKS module may have evolved at different evolutionary rates.

Characterization of the three operon gene structures in the *G. lozoyensis* genome suggests that operon structures might be common in the Fungi. Based on the prerequisites for an operon, such as two genes that are transcribed in the same direction with a predicted upstream promoter and an intergenic region of $\leq$700 bp that lacks predicted promoter sequences and polyadenylation signal sequences (13), we predicted the potential operon structures of three fungal genomes *in silico*. Genome and transcriptome data analysis revealed operonlike gene structures in 3.8%, 1.6%, and 3.7% of *A. nidulans*, *Glomerella graminicola*, and *Drechslerella stenobrocha* genes, respectively (see Table S2 in the supplemental material). If validated experimentally, this analysis suggests that operonlike gene structures may occur widely in the Fungi.

Xenolozoyenone, the novel secondary metabolite of GLPKS3-GLNRPS7 expressed in *A. nidulans*, is a typical fungal PKS-NRPS hybrid product. It belongs to the tetramic acid class of metabolites, such as ravenic acid (15), militarinone C (16), and pretenellin A (17), which are common in fungi. Its structure is similar to those of other tetramic acids but not identical due to having a different polyketide side chain and substituents on the 2,4-pyrrolidinedione ring. Moreover, the fungal hybrid polyketide-nonribosomal peptide molecules are generally assembled by a PKS-NRPS hybrid protein (18), including pyranonigrin E, biosynthesized by PynA in *Aspergillus niger*, which is one of the GLPKS3 KS domain orthologs showing a domain organization similar to that of GLPKS3-GLNPRS7 (19). However, GLPKS3-GLNRPS7 is an enzyme complex of two proteins, GLPKS3 and GLNRPS7. The PKS portion is responsible for the formation of the linear polyketide chain, and the NRPS portion for the exten-

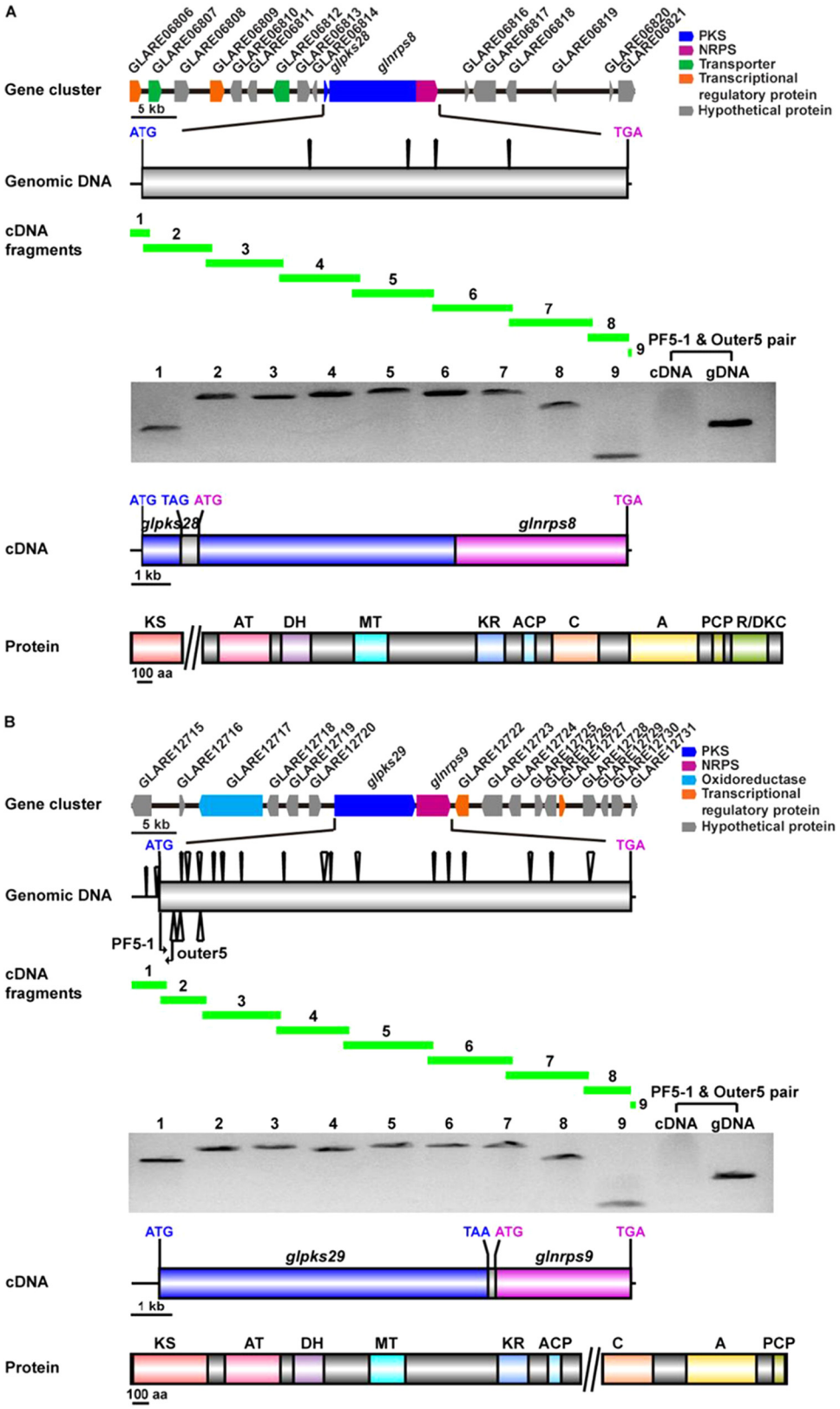

FIG 6 Genomic characterization of the *glpks28-glnrps8* and *glpks29-glnrps9* gene clusters. (A) *glpks28-glnrps8* gene clusters. The *glpks28-glnrps8* locus is located in a 54.4-kb gene cluster that includes two putative transporter genes and two putative transcriptional regulatory protein genes. (B) *glpks29-glnrps9* gene clusters. The *glpks29-glnrps9* locus is located in a 55.0-kb gene cluster that includes one putative oxidoreductase gene and two putative transcriptional regulatory protein genes. Vertical pointers indicate intron positions. cDNA fragments 1 and 9 were obtained by RACE-PCR, and fragments 2 to 8 were obtained by RT-PCR. The

(Continued)

sion of the polyketide chain with an activated glycine. The linear polyketide-nonribosomal peptide chain undergoes a Dieckmann cyclization reaction that results in the formation of compound 1a (17). Finally, compound 1a could be modified by tailoring enzymes (e.g., cytochrome P450) of *A. nidulans* to yield compound 1 (Fig. 4C). As the NRPS portion of GLPKS3-GLNRPS7 is located in two individual proteins, GLPKS3 (containing a condensation domain [C] and a partial A domain proximal to the C terminus) and GLNRPS7 (containing the complementary partial A domain, a peptidyl carrier protein [PCP], and a Dieckmann cyclase domain [DKC]), we presume that GLPKS3 and GLNRPS7 are translated in a 1:1 ratio to assemble the optimal number of functional enzyme complexes, and the operon genomic organization helps maintain the stoichiometric ratio (20).

The identification of operons in the Fungi not only reveals the complexity of gene transcription and regulation in eukaryotic organisms but also facilitates future studies concerning the origin and evolution of fungal genes that encode antibiotics and other secondary metabolites. These findings may have important practical implications in the construction of synthetic gene clusters for antibiotic discovery.

MATERIALS AND METHODS

DNA and RNA procedures. Fungal genomic DNA was extracted as described previously (10). Plasmids and PCR products were purified using the E.Z.N.A. gel extraction kit (Omega Bio-Tek, Norcross, GA). Phusion high-fidelity DNA polymerase (New England Biolabs, Ipswich, Massachusetts) and EasyTaq DNA polymerase (TransGen Biotech, Beijing, China) were used for PCRs. Sequences were determined by SinoGenoMax Co., Ltd. (Beijing, China). Restriction endonucleases and DNA-modifying enzymes were from New England Biolabs (Ipswich, Massachusetts).

RNA was extracted with the Qiagen RNeasy minikit, and carryover DNA was removed by DNase I digestion (Qiagen, Valencia, CA). cDNA fragments were synthesized using the TransScript II first-strand cDNA synthesis supermix (TransGen Biotech, Beijing, China). Introns were identified by comparing genomic and cDNA sequences. The 5′ ends of the mRNA and the poly(A) attachment sites were mapped by 5′- and 3′-RACE–PCR (FirstChoice RLM-RACE kit; Ambion, Austin, Texas).

All primers used in the study are listed in Table S3 in the supplemental material.

Genetic manipulation of *glpks3-glnrps7* in *G. lozoyensis*. To detect the expression of GLPKS3, a genomic fragment extending from the promoter to the region annotated as the KS domain of *glpks3*, the GFP reporter gene, and the *A. nidulans trpC* terminator were amplified and inserted into pAg1-H3 to give the expression vector pAg1-H3L-KS2-TtrpC (see Text S1 and Fig. S6 in the supplemental material).

To disrupt *glpks3*, the 3′ regions of *glpks3* and *glnrps7* were amplified and inserted into pAg1-H3 to give pAg1-H3-glpks3-glnrps7, containing the gene replacement cassette (see Fig. S1 in the supplemental material).

Conidia of *G. lozoyensis* were prepared and transformed as previously described (10). After isolation of single spores from the transformants, the positive transformants of Δ*glpks3-glnrps7* were determined by PCR using two primer pairs (GLPKS3-F2/999′ and GLPKS3-F1/T). Transformants positive for KS::GFP were determined by PCR using two primer pairs (GFP-F/GFP-R and GFP-F/N).

The wild-type strain of *G. lozoyensis*, the KS::GFP transformant, and the Δ*glpks3-glnrps7* mutant were grown in LYCP-5 medium to prepare seed cultures (10). The seed cultures were then inoculated into H medium and grown for 14 days (21). For the wild-type strain of *G. lozoyensis*, the culture was divided into two parts. One part was examined by RT-PCR using two primer pairs (Y/Z and I/999′) and Western blotting using the CAS2 antibody. The other part and the culture of the Δ*glpks3-glnrps7* strain were analyzed by HPLC-MS following the same procedures used for *A. nidulans*. For the KS::GFP transformant, mycelia were collected and observed by fluorescence microscopy.

Heterologous expression of *glpks3-glnrps7* in *A. nidulans*. The *glpks3-glnrps7* gene was amplified from genomic DNA of *G. lozoyensis*, and the constitutive *A. nidulans trpC* promoter was amplified from pAg1-H3 (22). After restriction enzyme digestion and ligation, *glpks3-glnrps7* under the control of the constitutive *A. nidulans trpC* promoter was constructed and inserted into pAg1-H3 to give the expression vector pAg1-PtrpC-glpks3-glnrps7 and control vector pAg1-glpks3-glnrps7 (see Text S1 and Fig. S6 in the supplemental material). Protoplast preparation and transformation were performed as previously described (23). One microgram each of pRG3-AMA1 and SbfI-digested pAg1-PtrpC-glpks3-glnrps7 or pAg1-glpks3-glnrps7 were cotransformed into *A. nidulans* A773 (*pyrG89 wA3 pyroA4 veA1*). Positive *glpks3-glnrps7*-containing transformants were identified by PCR using five primer sets (Pc800F/P22′, A/B, U/F, and M/D for PtrpC-glpks3-glnrps7 transformants and K/D for *glpks3-glnrps7* transformants) and DNA sequencing.

The positive transformants and parent strain of *A. nidulans* were incubated on MAG with appropriate supplements (23). Conidia were harvested and adjusted to 10^8 conidia ml^{-1}. Amounts of about 2 ml of spore suspension were inoculated into MMV (1% glucose, 0.001% thiamine, nitrate salts, trace elements, and vitamins, pH 6.5) with appropriate supplements (23). The cultures were shaken at 200 rpm in the dark for 7 days. RNA preparations from the cultures were examined for GLNRPS7 transcription and translation by RT-PCR using two primer pairs (Y/Z, and I/999′) for all strains of *A. nidulans*, five primer pairs (Y/P30R-2, GLPKS3-F7/U-R, U/GLPKS3-F2-R, GLPKS3-F2/GLNRPS7-F-R, and I/GLNRPS7-R1) for PtrpC-glpks3-glnrps7 transformants, and protein preparations were examined by Western blotting using the polyclonal antibody CAS2.

Protein analysis. Protein extraction, quantification, and Western blotting were performed as previously described (24). In order to detect GLNRPS7, we prepared a polyclonal antibody, CAS2, by immunizing a rabbit (Epitomics, Hangzhou, Zhejiang, China) with the peptide antigen PSSKSNEPA, which was predicted as an immune epitope against the predicted GLNRPS7 protein by the PeptideStructure program of the University of Wisconsin Genetics Computer Group Sequence Analysis Software Package. The IgG titer against the PSSKSNEPA peptide antigen reached 64,000. The polyclonal antibody CAS2 was purified by protein A column as described previously (25).

Fermentation and extraction of *A. nidulans* and transformants for HPLC-MS analysis. The transformants and parent strain of *A. nidulans* were cultured on MAG with appropriate supplements (23). Fifteen pieces (0.5 by 0.5 by 0.5 cm^3) of agar cultures were inoculated and incubated in 100 ml H medium with appropriate supplements at 30°C and 220 rpm for 7 days. The cultures were extracted with 100 ml of methyl ethyl ketone (MEK), and the organic phase was evaporated to dryness and redissolved in methanol (MeOH) to 30 $mg \cdot ml^{-1}$.

NMR and MS methods. All ^1H, ^{13}C, and two-dimensional (^1H–^1H correlation spectroscopy [COSY], heteronuclear single quantum coherence [HSQC] spectroscopy, and heteronuclear multiple-bond correlation [HMBC] spectroscopy) NMR spectra were acquired on a Bruker 600- or

Figure Legend Continued

whole transcript was assembled based on the overlapping cDNA fragments. Primer set PF5-1 and Outer5 was designed to amplify a 404-bp sequence from genomic DNA templates but not from cDNA templates. KS, ketosynthase domain; AT, acyltransferase domain; DH, dehydratase domain; MT, methyltransferase; KR, β-ketoacylreductase domain; ACP, acyl carrier protein; C, condensation domain; A, adenylation domain; PCP, peptidyl carrier protein; R, reductive domain; DKC, Dieckmann cyclase domain.

500-MHz spectrometer equipped with a 5-mm triple-resonance cryoprobe at 298 K. Residual solvent signals were used as references (acetone-d_6, δ_H 2.05/δ_C 29.8, 206.0, and $CDCl_3$, δ_H 7.26/δ_C 77.0). HPLC-MS spectra were obtained on an Agilent 6120 quadrupole mass spectrometer using a positive ESI source. Amounts of 10 μl of the test samples were injected for HPLC-MS analysis (Agilent Zorbax Eclipse plus C_{18} reverse-phase column, 5 μm, 4.6 by 150 mm, 10% to 90% CH_3CN in H_2O with 01% formic acid for 30 min, 1 ml·min^{-1}). HRESIMS data were obtained using the Agilent 6520 quadrupole time of flight (Q-TOF) LC-MS instrument equipped with an electrospray ionization (ESI) source.

Isolation of xenolozoyenone. The scaled-up fermentation culture (*PtrpC-glpks3-glnrps7* transformant of *A. nidulans* in 2.0 liters of H medium) was extracted repeatedly with MEK (3 times per 2.0 liters), and the organic solvent was vacuum evaporated to dryness to obtain the crude extract (1.2 g), which was fractionated by using a Sephadex LH-20 column with 1:1 CH_2Cl_2-MeOH as eluents. The resulting subfractions were combined and further purified by semipreparative reverse-phase HPLC (Agilent Zorbax SB-C_{18} column, 5 μm, 9.4 by 250 mm, 65% CH_3CN in H_2O with 01% formic acid for 25 min, 2 ml·min^{-1}) to obtain xenolozoyenone (compound 1) as follows: 4.5 mg, t_R 17.25 min (MeOH); UV (MeOH) λ_{max} (log ε) 217 (4.46), 256 (4.10), 390 (4.86) nm; ^{1}H NMR (acetone-d_6, 600 MHz) δ 7.73 (1H, brs, OH-1 or OH-6), 7.50 (1H, dd, J = 15.0, 11.4 Hz, H-8), 7.12 (1H, d, J = 15.0 Hz, H-7), 6.82 (1H, dd, J = 15.0, 10.8 Hz, H-10), 6.50 (1H, dd, J = 15.0, 11.4 Hz, H-9), 6.32 (1H, dd, J = 15.0, 10.8 Hz, H-11), 6.10 (1H, dt, J = 15.0, 7.2 Hz, H-12), 3.81 (2H, m, H_2-5), 2.19 (2H, m, H_2-13), 1.44 (2H, m, H_2-14), 1.31 (6H, m, H_2-15, H_2-16, and H_2-17), 0.89 (3H, t, J = 6.6, H_3-18); ^{13}C NMR ($CDCl_3$, 125 MHz) δ 193.4 (C-4), 176.1 (C-2), 175.0 (C-6), 146.3 (C-8), 144.6 (C-10), 143.4 (C-12), 130.2 (C-11), 128.8 (C-9), 120.1 (C-7), 99.7 (C-3), 62.9 (C-5), 33.2 (C-13), 31.4 (C-15), 29.7 (C-17), 28.6 (C-14), 22.5 (C-16), 14.0 (C-18); HRESIMS m/z $[M+H]^+$ 306.1708 (calculated for $C_{17}H_{24}NO_4$, 306.1700).

Phylogenetic analysis of PKSs and PKS-NRPSs. To test whether the *glpks3-glnrps7* locus is a product of horizontal gene transfer from bacteria, we built a phylogenetic tree based on the KS domain of GLPKS3-GLNRPS7, its orthologs in other fungi (coverage, ≥95%, and identity, ≥64%), GLPKS28-GLNRPS8, GLPKS29-GLNRPS9, 30 characterized fungal PKSs, and 9 characterized bacterial PKSs (see Table S1 in the supplemental material). All domains were identified by using the program anti-SMASH (26) or manually in multiple alignments. The amino acid sequences of the KS domains were aligned with MUSCLE and analyzed phylogenetically with MEGA 6.0 by the maximum-likelihood (ML) algorithm using an LG+F+G+I model selected by Prot-test and a 1,000-replication bootstrap test (27, 28). The R/DKC domains were also aligned with MUSCLE. To test whether the *glpks3-glnrps7* locus might be of single or multiple evolutionary origins, we constructed two phylogenetic trees, one based on the KS domain of GLPKS3, all orthologs in other fungi (coverage, ≥95%, and identity, ≥55%), and 3 characterized fungal PKS-NRPSs and the other based on the A domain of GLPKS3, all orthologs in other fungi (coverage, ≥95%, and identity, ≥40%), and 3 characterized fungal PKS-NRPSs. The KS and A domains were identified by the program anti-SMASH (26) or manually in multiple alignments. The amino acid sequences of these domains were aligned with MUSCLE and analyzed with RAxML BlackBox by the ML method under a WAG (Whelan And Goldman) model with or without a constraint from the reference fungal taxonomy (27, 29). The A domain tree was also analyzed with RAxML BlackBox by the ML method under the WAG model with a constraint from the KS domain tree.

Alternative hypotheses based on the tree topologies under the null hypothesis that all topologies are equally good explanations of the data were tested with the Shimodaira-Hasegawa test (30), weighted Shimodaira-Hasegawa test, and approximately unbiased test (31), as implemented in TREEFINDER (32).

Bioinformatics. Putative promoter sequences in the intergenic region of *glpks3-glnrps7* and 1 kb upstream from the genes in other gene pairs were investigated with Promoter 2.0 Prediction Server (33). The polyadenylation signals (AATAAA and ATTAAA) in the intergenic region of *glpks3-glnrps7* and other gene pairs were identified by using the nucleotide sequence pattern search tool Fuzznuc in EMBOSS (http://emboss.bioinformatics.nl/cgi-bin/emboss/fuzznuc). To identify potential polycistrons, genome and transcriptome data of *A. nidulans* (GenBank accession numbers PRJNA13961 and PRJNA182228), *G. graminicola* (GenBank accession numbers PRJNA37879 and PRJNA151285), and *D. stenobrocha* (GenBank accession numbers PRJNA67941 and PRJNA236481) were downloaded from NCBI (http://www.ncbi.nlm.nih.gov/). The genomic loci were annotated by InterProScan, KEGG, FunCat, and UniProt-GO (34–37) to rule out the possibility of pseudogenes. The RNA-Seq data were assembled as described previously and then mapped and aligned to the genome (38, 39).

The mRNA sequences of *glpks3-glnrps7*, *glpks28-glnrps8*, and *glpks29-glnrps9* have been deposited at GenBank under the accession numbers KM603664, KM603665, and KM603666, respectively.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.00703-15/-/DCSupplemental.

Text S1, DOC file, 0.04 MB.
Figure S1, PDF file, 0.1 MB.
Figure S2, PDF file, 0.2 MB.
Figure S3, PDF file, 0.2 MB.
Figure S4, PDF file, 0.3 MB.
Figure S5, PDF file, 0.2 MB.
Figure S6, PDF file, 0.5 MB.
Table S1, DOC file, 0.2 MB.
Table S2, DOC file, 0.03 MB.
Table S3, DOC file, 0.1 MB.

ACKNOWLEDGMENTS

This work was supported by National Natural Science Foundation of China (NSFC) grants 31328001 to Z.A. and 31200055 to Q.Y., Welch Foundation grant AU00024 to Z.A., a University of Texas System Star Award to Z.A., and a grant from the University of Texas Health Science Center at Houston faculty startup fund to G.F.B.

REFERENCES

1. **Field B, Fiston-Lavier AS, Kemen A, Geisler K, Quesneville H, Osbourn AE.** 2011. Formation of plant metabolic gene clusters within dynamic chromosomal regions. Proc Natl Acad Sci U S A **108:**16116–16121. http://dx.doi.org/10.1073/pnas.1109273108.
2. **Blumenthal T.** 2004. Operons in eukaryotes. Brief Funct Genomic Proteomic **3:**199–211. http://dx.doi.org/10.1093/bfgp/3.3.199.
3. **Spieth J, Brooke G, Kuersten S, Lea K, Blumenthal T.** 1993. Operons in *C. elegans*: polycistronic mRNA precursors are processed by *trans*-splicing of SL2 to downstream coding regions. Cell **73:**521–532. http://dx.doi.org/10.1016/0092-8674(93)90139-H.
4. **Evans D, Zorio D, MacMorris M, Winter CE, Lea K, Blumenthal T.** 1997. Operons and SL2 *trans*-splicing exist in nematodes outside the genus *Caenorhabditis*. Proc Natl Acad Sci U S A **94:**9751–9756. http://dx.doi.org/10.1073/pnas.94.18.9751.
5. **Satou Y, Hamaguchi M, Takeuchi K, Hastings KE, Satoh N.** 2006. Genomic overview of mRNA 5′-leader *trans*-splicing in the ascidian *Ciona intestinalis*. Nucleic Acids Res **34:**3378–3388. http://dx.doi.org/10.1093/nar/gkl418.
6. **Blumenthal T, Evans D, Link CD, Guffanti A, Lawson D, Thierry-Mieg J, Thierry-Mieg D, Chiu WL, Duke K, Kiraly M, Kim SK.** 2002. A global analysis of *Caenorhabditis elegans* operons. Nature **417:**851–854. http://dx.doi.org/10.1038/nature00831.
7. **Gray TA, Saitoh S, Nicholls RD.** 1999. An imprinted, mammalian bicistronic transcript encodes two independent proteins. Proc Natl Acad Sci U S A **96:**5616–5621. http://dx.doi.org/10.1073/pnas.96.10.5616.
8. **García-Ríos M, Fujita T, LaRosa PC, Locy RD, Clithero JM, Bressan RA, Csonka LN.** 1997. Cloning of a polycistronic cDNA from tomato encoding γ-glutamyl kinase and γ-glutamyl phosphate reductase. Proc

Natl Acad Sci U S A **94:**8249–8254. http://dx.doi.org/10.1073/pnas.94.15.8249.
9. **Slone J, Daniels J, Amrein H.** 2007. Sugar receptors in *Drosophila*. Curr Biol **17:**1809–1816. http://dx.doi.org/10.1016/j.cub.2007.09.027.
10. **Chen L, Yue Q, Zhang X, Xiang M, Wang C, Li S, Che Y, Ortiz-López FJ, Bills GF, Liu X, An Z.** 2013. Genomics-driven discovery of the pneumocandin biosynthetic gene cluster in the fungus *Glarea lozoyensis*. BMC Genomics **14:**339. http://dx.doi.org/10.1186/1471-2164-14-339.
11. **Proudfoot NJ.** 2011. Ending the message: poly(A) signals then and now. Genes Dev **25:**1770–1782. http://dx.doi.org/10.1101/gad.17268411.
12. **Chung KR, Daub ME, Ehrenshaft M.** 2003. Expression of the cercosporin toxin resistance gene (*CRG1*) as a dicistronic mRNA in the filamentous fungus *Cercospora nicotianae*. Curr Genet **43:**415–424. http://dx.doi.org/10.1007/s00294-003-0414-3.
13. **Ben-Shahar Y, Nannapaneni K, Casavant TL, Scheetz TE, Welsh MJ.** 2007. Eukaryotic operon-like transcription of functionally related genes in *Drosophila*. Proc Natl Acad Sci U S A **104:**222–227. http://dx.doi.org/10.1073/pnas.0609683104.
14. **Andrews J, Smith M, Merakovsky J, Coulson M, Hannan F, Kelly LE.** 1996. The stoned locus of *Drosophila melanogaster* produces a dicistronic transcript and encodes two distinct polypeptides. Genetics **143:**1699–1711.
15. **Michael AP, Grace EJ, Kotiw M, Barrow RA.** 2002. Ravenic acid, a new tetramic acid isolated from a cultured microfungus, *Penicillium* sp. J Nat Prod **65:**1360–1362. http://dx.doi.org/10.1021/np0200358.
16. **Schmidt K, Riese U, Li Z, Hamburger M.** 2003. Novel tetramic acids and pyridone alkaloids, militarinones B, C, and D, from the insect pathogenic fungus *Paecilomyces militaris*. J Nat Prod **66:**378–383. http://dx.doi.org/10.1021/np020430y.
17. **Halo LM, Marshall JW, Yakasai AA, Song Z, Butts CP, Crump MP, Heneghan M, Bailey AM, Simpson TJ, Lazarus CM, Cox RJ.** 2008. Authentic heterologous expression of the tenellin iterative polyketide synthase nonribosomal peptide synthetase requires coexpression with an enoyl reductase. Chembiochem **9:**585–594. http://dx.doi.org/10.1002/cbic.200700390.
18. **Boettger D, Hertweck C.** 2013. Molecular diversity sculpted by fungal PKS-NRPS hybrids. Chembiochem **14:**28–42. http://dx.doi.org/10.1002/cbic.201200624.
19. **Awakawa T, Yang XL, Wakimoto T, Abe I.** 2013. Pyranonigrin E: a PKS-NRPS hybrid metabolite from *Aspergillus niger* identified by genome mining. Chembiochem **14:**2095–2099. http://dx.doi.org/10.1002/cbic.201300430.
20. **Bratlie MS, Johansen J, Drabløs F.** 2010. Relationship between operon preference and functional properties of persistent genes in bacterial genomes. BMC Genomics **11:**71. http://dx.doi.org/10.1186/1471-2164-11-71.
21. **Tkacz JS, Giacobbe RA, Monaghan RL.** 1993. Improvement in the titer of echinocandin-type antibiotics: a magnesium-limited medium supporting the biphasic production of pneumocandins A_0 and B_0. J Ind Microbiol **11:**95–103. http://dx.doi.org/10.1007/BF01583681.
22. **Zhang A, Lu P, Dahl-Roshak AM, Paress PS, Kennedy S, Tkacz JS, An Z.** 2003. Efficient disruption of a polyketide synthase gene (*pks1*) required for melanin synthesis through *Agrobacterium*-mediated transformation of *Glarea lozoyensis*. Mol Genet Genomics **268:**645–655. http://dx.doi.org/10.1007/s00438-002-0780-4.
23. **Li S, Du L, Yuen G, Harris SD.** 2006. Distinct ceramide synthases regulate polarized growth in the filamentous fungus *Aspergillus nidulans*. Mol Biol Cell **17:**1218–1227. http://dx.doi.org/10.1091/mbc.E05-06-0533.
24. **Garceau NY, Liu Y, Loros JJ, Dunlap JC.** 1997. Alternative initiation of translation and time-specific phosphorylation yield multiple forms of the essential clock protein FREQUENCY. Cell **89:**469–476. http://dx.doi.org/10.1016/S0092-8674(00)80227-5.
25. **Yu Y, Lee P, Ke Y, Zhang Y, Yu Q, Lee J, Li M, Song J, Chen J, Dai J, Do Couto FJ, An Z, Zhu W, Yu GL.** 2010. A humanized anti-VEGF rabbit monoclonal antibody inhibits angiogenesis and blocks tumor growth in xenograft models. PLoS One **5:**e9072. http://dx.doi.org/10.1371/journal.pone.0009072.
26. **Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R.** 2011. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res **39:**W339–W346. http://dx.doi.org/10.1093/nar/gkr466.
27. **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S.** 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol **30:**2725–2729. http://dx.doi.org/10.1093/molbev/mst197.
28. **Abascal F, Zardoya R, Posada D.** 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics **21:**2104–2105. http://dx.doi.org/10.1093/bioinformatics/bti263.
29. **Stamatakis A, Hoover P, Rougemont J.** 2008. A rapid bootstrap algorithm for the RAxML web servers. Syst Biol **57:**758–771. http://dx.doi.org/10.1080/10635150802429642.
30. **Shimodaira H, Hasegawa M.** 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol Biol Evol **16:**1114–1116. http://dx.doi.org/10.1093/oxfordjournals.molbev.a026201.
31. **Shimodaira H.** 2002. An approximately unbiased test of phylogenetic tree selection. Syst Biol **51:**492–508. http://dx.doi.org/10.1080/10635150290069913.
32. **Jobb G, von Haeseler A, Strimmer K.** 2004. TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. BMC Evol Biol **4:**18. http://dx.doi.org/10.1186/1471-2148-4-18.
33. **Knudsen S.** 1999. Promoter2.0: for the recognition of PolII promoter sequences. Bioinformatics **15:**356–361.
34. **Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R.** 2005. InterProScan: protein domains identifier. Nucleic Acids Res **33:**W116–W120. http://dx.doi.org/10.1093/nar/gki442.
35. **Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M.** 2012. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res **40:**D109–D114. http://dx.doi.org/10.1093/nar/gkr988.
36. **Ruepp A, Zollner A, Maier D, Albermann K, Hani J, Mokrejs M, Tetko I, Güldener U, Mannhaupt G, Münsterkötter M, Mewes HW.** 2004. The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. Nucleic Acids Res **32:**5539–5545. http://dx.doi.org/10.1093/nar/gkh894.
37. **Dimmer EC, Huntley RP, Alam-Faruque Y, Sawford T, O'Donovan C, Martin MJ, Bely B, Browne P, Mun Chan W, Eberhardt R, Gardner M, Laiho K, Legge D, Magrane M, Pichler K, Poggioli D, Sehra H, Auchincloss A, Axelsen K, Blatter MC, Boutet E, Braconi-Quintaje S, Breuza L, Bridge A, Coudert E, Estreicher A, Famiglietti L, Ferro-Rojas S, Feuermann M, Gos A, Gruaz-Gumowski N, Hinz U, Hulo C, James J, Jimenez S, Jungo F, Keller G, Lemercier P, Lieberherr D, Masson P, Moinat M, Pedruzzi I, Poux S, Rivoire C, Roechert B, Schneider M, Stutz A, Sundaram S, Tognolli M, Bougueleret L.** 2012. The UniProt-GO annotation database in 2011. Nucleic Acids Res **40:**D565–D570. http://dx.doi.org/10.1093/nar/gkr1048.
38. **Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A.** 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol **29:**644–652. http://dx.doi.org/10.1038/nbt.1883.
39. **Wu TD, Watanabe CK.** 2005. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics **21:**1859–1875. http://dx.doi.org/10.1093/bioinformatics/bti310.