Alternative transcription start sites of the enolase-encoding gene enoA are stringently used in glycolytic/gluconeogenic conditions in Aspergillus oryzae

doi: 10.1007/s00294-020-01053-3

Title: Alternative transcription start sites of the enolase-encoding gene *enoA* are stringently used in glycolytic/gluconeogenic conditions in *Aspergillus oryzae*

Author:

Taishi Inoue, Hiroki Toji, Mizuki Tanaka[‡], Mitsuru Takama, Sachiko Hasegawa-Shiro[†], Yuichi Yamaki, Takahiro Shintani, Katsuya Gomi*

Laboratory of Bioindustrial Genomics, Department of Bioindustrial Informatics and Genomics, Graduate School of Agricultural Science, Tohoku University, 468-1 Aoba, Aramaki, Aoba-ku, Sendai 980-8572, Japan

*Corresponding author:

Tel./Fax.: +81-22-757-4489;

E-mail address: katsuya.gomi.a6@tohoku.ac.jp (K. Gomi)

ORCID iD: 0000-0003-3463-8072

[‡]Present address: Biomolecular Engineering Laboratory, School of Food and Nutritional Science, University of Shizuoka, 52-1 Yada, Suruga-ku, Shizuoka 422-8526, Japan

[†]Present address: Department of Food and Brewing Technology, Yamagata Research Institute of Technology, 2-2-1 Shoei, Yamagata 990-2473, Japan

Acknowledgements:

This work was supported by the Division for Interdisciplinary Advanced Research and Education (DIARE) Tohoku University. We would like to thank Editage

25 (www.editage.com) for English language editing.

26

Abstract:

Gene expression by using alternative transcription start sites (TSSs) is an important transcriptional regulatory mechanism for environmental responses in eukaryotes. Here, we identify two alternative TSSs in the enolase-encoding gene (*enoA*) in *Aspergillus oryzae*, an industrially important filamentous fungus. TSS use in *enoA* is strictly dependent on the difference in glycolytic and gluconeogenic carbon sources. Transcription from the upstream TSS (uTSS) or downstream TSS (dTSS) predominantly occurs under gluconeogenic or glycolytic conditions, respectively. In addition to *enoA*, most glycolytic genes involved in reversible reactions possess alternative TSSs. The *fbaA* gene, which encodes fructose-bisphosphate aldolase, also shows stringent alternative TSS selection, similar to *enoA*. Alignment of promoter sequences of enolase-encoding genes in *Aspergillus* predicted two conserved regions that contain a putative *cis*-element required for *enoA* transcription from each TSS. However, uTSS-mediated transcription of the *acuN* gene, an *enoA* ortholog in *Aspergillus nidulans*, is not strictly dependent on carbon source, unlike *enoA*. Furthermore, *enoA* transcript levels in glycolytic conditions are higher than in gluconeogenic conditions. Conversely, *acuN* is more highly transcribed in gluconeogenic conditions. This suggests that the stringent usage of alternative TSSs and higher transcription in glycolytic conditions in *enoA* may reflect that the *A. oryzae* evolutionary genetic background was domesticated by exclusive growth in starch-rich environments. These findings provide novel insights into the complexity and diversity of transcriptional regulation of glycolytic/gluconeogenic genes among *Aspergilli*.

Keywords: *Aspergillus oryzae*; alternative transcription start site; glycolytic gene; enolase; gluconeogenesis; AcuK/AcuM

Introduction:

Fungi display versatile metabolisms of carbon sources. Carbon metabolism plays a role in their pathogenicity and chemical production, which are required for growth. Glycolysis and gluconeogenesis are primary metabolic pathways of carbon sources. Glycolysis is involved in glucose catabolism accompanied by substrate level phosphorylation, while gluconeogenesis is involved in glucose anabolism required for providing start materials for synthesizing cellular components such as nucleic acids and sugar chains. Therefore, elucidating regulatory mechanisms of these metabolic pathways is fundamentally important for understanding characteristic metabolic features and survival strategies of fungal species.

Aspergillus oryzae is among the most important filamentous fungi used in fermentation industries. It has been extensively used to produce traditional Japanese fermented beverages and foods, such as sake (rice wine), shoyu (soy sauce), and miso (soybean paste), for over a thousand years (Machida et al., 2008). *A. oryzae* is also a promising host to produce heterologous recombinant proteins for industrial use because of its ability to secrete large amounts of hydrolytic enzymes (Oda et al., 2006; Tanaka and Gomi, 2013). In addition, its safety is supported by extensive use in food production (Barbesgaard et al., 1992; Machida et al., 2008). Furthermore, *A. oryzae* can produce organic acids (Brown et al., 2013; Wakai et al., 2014; Yang et al., 2017) and heterologous secondary metabolites with medical properties (Sakai et al., 2012; Tagami et al., 2013; Asai et al., 2015; Liu et al., 2015; Fujii et al., 2016; Yoshimi et al., 2018). Thus, interest in the molecular details of *A. oryzae* is increasing.

In *A. oryzae*, most glycolytic genes are strongly expressed in the presence of fermentable carbon sources like glucose (Nakajima et al., 2000; Maeda et al., 2004). In

particular, the *enoA* gene encodes enolase (2-phospho-D-glycerate hydrolase, EC 4.2.1.11) and is among the most highly expressed glycolytic genes in *A. oryzae* (Machida et al., 1996). Previous studies suggest that the *enoA* transcript level is comparable to the Taka-amylase A (TAA) gene (*amyB*) that is very strongly expressed in *A. oryzae*. The *enoA* transcript comprises approximately 3% (w/w) of total mRNA (Machida et al., 1996). Interestingly, high enolase gene expression was also reported in *Saccharomyces cerevisiae*, an important microorganism in fermentation industries (Holland and Holland, 1978). Thus, high-level enolase gene expression might be fundamentally important for both *A. oryzae* and *S. cerevisiae*. Additionally, promoters of glycolytic genes may be useful tools for the high-level production of recombinant proteins in *A. oryzae*. Indeed, *enoA* promoter improvement has been attempted for industrial use (Tsuboi et al., 2005). Therefore, understanding the molecular regulatory mechanisms of glycolytic gene expression in *A. oryzae* is important for both biological and biotechnological aspects. However, despite their significance, most transcriptional machineries involved in glycolytic gene expression remain unclear.

In *A. oryzae*, primer extension analysis indicates that the *enoA* transcription start sites are located at −44, −37, −31 and −17 base pairs upstream of the start codon (+1) when cultured with glucose (Machida et al., 1996). Deletion analysis of the *enoA* promoter showed that the deletion of a 104 bp region between −224 and −121 results in loss of promoter activity in the presence of glucose (Toda et al., 2001). Furthermore, electrophoretic gel mobility shift assay (EMSA) using whole cell extracts suggested that an unidentified regulator protein binds to the 15-bp region between −195 and −181 for high *enoA* expression (Toda et al., 2001). Conversely, a translocation mutation in *Aspergillus nidulans*, with a break point at −220 in the enolase-encoding gene (*acuN*)

results in the *acuN356* mutant strain being unable to utilize acetate (Armitt et al., 1976; Hynes et al., 2007). Intriguingly, the break point was located in a large intron between −394 and −10 in the 5′ untranslated region (5′ UTR). This mutation results in loss of *acuN* expression in the presence of non-fermentable carbon sources such as acetate and ethanol, but not of fermentable carbon sources, such as glucose and fructose (Hynes et al., 2007). In addition, *acuN* expression in cultures with non-fermentable carbon sources is regulated by the two transcription factors, AcuK and AcuM, which are involved in the regulation of gluconeogenesis (Hynes et al., 2007; Suzuki et al., 2012). These findings suggest that *enoA*/*acuN* expression is regulated by distinct mechanisms under culture conditions associated with glycolysis or gluconeogenesis, but those mechanisms remain unclear.

In this study, we investigated the molecular details of *enoA*/*acuN* expression mechanisms underlying the usage pattern of transcription start sites (TSSs). We demonstrate that the *A. oryzae enoA* gene has two TSSs, upstream TSS (uTSS) and downstream TSS (dTSS), which were strictly used to respond to different carbon sources associated with glycolysis or gluconeogenesis. In addition, we identified two highly conserved sequences are present in enolase-encoding gene promoters in *Aspergillus* fungi that contain *cis*-enhancer elements required for *enoA* transcription from each TSS. Interestingly, the induction of the two TSSs and resulting transcript levels between *enoA* and *acuN* differ depending on the carbon source species. Our findings provide novel insights on complex and diverse gene regulatory mechanisms involved in *Aspergillus* primary metabolic pathways.

Materials and Methods:

123 **Strains and Media**

124 *Aspergillus oryzae* RIB40 (Machida et al. 2005; National Research Institute of
125 Brewing Stock Culture, Higashi-Hiroshima, Japan) was used as the wild-type strain for
126 northern blot analysis, 5′ serial analysis of gene expression (5′ SAGE), 5′ rapid
127 amplification of cDNA ends (5′ RACE), and quantitative reverse transcription-PCR (qRT-
128 PCR) analysis. *Aspergillus nidulans* FGSC A4 strain was also used for 5′ RACE, qRT-
129 PCR, and northern blot analyses. For the construction of strains for β-glucuronidase
130 (GUS) reporter assays, *A. oryzae* niaD300 strain (*niaD*$^{-}$) derived from RIB40 (Minetoki
131 et al., 1996) was used as the transformation recipient strain. For *acuK* or *acuM* disruption,
132 the *A. oryzae* Δ*ligD*::*ptrA* strain (*niaD*$^{-}$, *sC*$^{-}$), derived from the Δ*ligD*::*sC* strain (Mizutani
133 et al., 2008), was used as the recipient strain. The *acuK* or *acuM* disruptant complemented
134 with *niaD* was used as the Δ*acuK* or Δ*acuM* strain. The Δ*ligD*::*sC* strain complemented
135 with *niaD* was used as a control strain for the *acuK* or *acuM* disruptant. *Escherichia coli*
136 DH5α (Hanahan et al., 1983) was used to construct and propagate plasmid DNAs for *A.*
137 *oryzae* transformation.

138 Medium containing 0.5% yeast extract, 1% peptone, and 1% glucose (YPD) was used
139 as complete culture medium. Wheat bran solid medium contained 2 g wheat bran, 0.08 g
140 $(NH_4)_2SO_4$, 0.03 g KH_2PO_4, 0.04 g maltose, and was moistened with 2 mL H_2O. Czapek–
141 Dox medium (0.6% $NaNO_3$, 0.05% KCl, 0.2% KH_2PO_4, 0.05% $MgSO_4$, and trace
142 amounts of $FeSO_4$, $ZnSO_4$, $CuSO_4$, $MnSO_4$, $Na_2B_4O_7$, and $(NH_4)_6Mo_7O_{24}$, and 2% carbon
143 source) was used as minimal medium (MM). To cultivate the *niaD*-deficient strains in
144 MM, 0.6% $NaNO_3$ was replaced with 0.5% $(NH_4)_2SO_4$ as the nitrogen source. To cultivate
145 the *sC*-deficient strains, 0.0003% (0.02 mM) methionine was added to MM. LB+amp *E.*
146 *coli* culture medium contained 1% tryptone, 0.5% yeast extract, 0.5% NaCl, and 0.005%

147 ampicillin.

148

149 **Total RNA preparation**

150 Total RNA samples from mycelia grown in submerged cultures were prepared as
151 follows: Harvested mycelia were washed with water. Excess liquid was removed with
152 blotting paper, and samples were immediately frozen in liquid nitrogen and stored at
153 −80□. Frozen mycelia were ground to fine powder using a mortar and pestle in liquid
154 nitrogen, then suspended in ISOGEN reagent (Nippon Gene, Tokyo, Japan). Total RNA
155 was purified according to the manufacturer's instructions. Total RNA samples from
156 mycelia grown in solid-state culture using wheat bran were prepared as previously
157 described (Akao et al., 2002).

158

159 **5′ cDNA ends analysis**

160 To deduce the putative TSSs of the genes of interest, sequences of expressed sequence
161 tags (ESTs) flanked by the start codon were retrieved from the *A. oryzae* EST database
162 (Akao et al., 2007; https://nribf21.nrib.go.jp/EST2/) and were compared to the genomic
163 sequence (Machida et al., 2005; http://www.aspgd.org/). 5′ SAGE analysis used total
164 RNAs prepared from mycelia grown in submerged and solid-state *A. oryzae* cultures. The
165 5′ SAGE analysis and obtained sequence tag annotation were performed by the Post
166 Genome Institute (Tokyo, Japan) as previously described (Hashimoto et al., 2004). 5′
167 RACE analysis was performed using total RNAs prepared from mycelia grown in
168 submerged *A. oryzae* or *A. nidulans* cultures using RNA ligase-mediated RACE (RLM-
169 RACE). A GeneRacer kit (Invitrogen, Carlsbad, CA) was used for 5′ RACE. Primers used
170 for 5′ RACE are listed in Supplementary Table 1.

171

172 **Quantitative RT-PCR (qRT-PCR) and RT-PCR analysis**

173 For RT-PCR analyses, 40−50 μg total RNA was treated with RNase-free recombinant
174 DNase I (Takara Bio Inc., Otsu, Japan). First-strand cDNA was synthesized using
175 PrimeScript □ RTase (Takara Bio Inc.) with oligo(dT) primers and 1 μg DNase-treated
176 total RNA. Synthesized cDNA was treated with RNase H (Invitrogen), diluted 1:10 in
177 sterile distilled water, and used as qRT-PCR and RT-PCR template. qRT-PCR analyses
178 used SYBR Green PCR Master Mix (Thermo Fisher Scientific, Waltham, MA) and a
179 StepOnePlus Real-Time PCR system (Life Technologies, Carlsbad, CA).

180 For the evaluation of TSS usage in the *A. oryzae enoA* gene or *A. nidulans acuN* gene,
181 primer sets were designed to discriminate transcripts derived from the two TSSs.
182 Difference in amplification efficiency was less than 5% between two primer sets to detect
183 each TSS-dependent transcript (data not shown). Primers designed to detect the CDS
184 were used for control signal. Ct values were calculated by setting the fluorescence
185 threshold to the ΔR_n value 1.0. The Ct value of control signal was subtracted from that of
186 each TSS-derived transcript. Finally, transcript values from each TSS relative to total
187 transcript levels were calculated from the subtracted amount of Ct values. Ratio of mRNA
188 expression level was calculated by the $-\Delta\Delta$Ct method (Livak and Schmittgen, 2001). The
189 histone *H4* gene was used as a reference gene. RT-PCR analysis was performed using *Ex-*
190 *Taq* polymerase (Takara Bio Inc.) followed by 2% agarose gel electrophoresis. The gel
191 was stained with ethidium bromide (EtBr) and the PCR products were detected using an
192 ultraviolet transilluminator. Primers used for qRT-PCR and RT-PCR are listed in
193 Supplemental Table 1.

194

Northern blot analysis

Approximately 20 μg total RNA was electrophoresed on a formaldehyde-denatured 1.0% agarose gel, stained with EtBr, and transferred onto a Hybond-N+ membrane (GE Healthcare, Buckinghamshire, UK) using the capillary transfer method with 3 M NaCl and 0.3 M sodium citrate (SSC) transfer buffer. Digoxigenin (DIG)-labeled DNA fragments were synthesized using a PCR DIG Probe Synthesize Kit (Roche Diagnostics, Tokyo, Japan). PCR was performed using *A. oryzae* or *A. nidulans* cDNA and the primer set enoA-NP_Fw + enoA-NP_Rv, acuN-NP_Fw + acuN-NP_Rv, and uidA-NP_Fw + uidA-NP_Rv to synthesize probes to detect *enoA*, *acuN*, and *uidA* transcripts, respectively. Hybridization and signal detection were performed according to manufacturer instructions (Roche Diagnostics). An ImageQuant LAS 4000 instrument (GE Healthcare) was used to detect EtBr-stained rRNA and transcript. Signal intensity was quantified using ImageJ software (https://imagej.net/ImageJ).

Plasmid DNA construction

Primers and plasmid DNA used in this study are listed in Supplementary Table 1 and Supplementary Table 2, respectively. Plasmid DNA was constructed and used for promoter activity assays using the *E. coli* β-glucuronidase gene (*uidA*). The DNA fragment 1,182 bp upstream of the *enoA* start codon was amplified by PCR using *A. oryzae* genomic DNA and PenoA_Fw + PenoA_Rv primers. The amplified DNA fragment was digested with *Pst*I and *Xho*I and inserted into *Pst*I/*Sal*I-digested pNGAG1 (Fujioka et al., 2007), yielding pNPenoAGUS. The pNGAG1 plasmid was constructed by inserting the *glaA* promoter into *Pst*I/*Sal*I-digested pNAGT4 (Minetoki et al., 1996). pNAGT4 was used to produce the negative control strain in GUS reporter assay.

Plasmid DNA used to express *uidA* under control of the *enoA* promoter including a 103 bp deletion between −121 nt and −224 (PenoAΔ−121 to −224) was constructed as follows: PenoAΔ−121 to −224 was amplified by fusion PCR. DNA fragments between −225 and −1182 and between −1 and −120 were amplified by PCR using *A. oryzae* genomic DNA, PenoA_Fw + PenoAΔ−121 to −224_Rv, and PenoAΔ−121 to −224_Fw + PenoA_Rv primers, respectively. The two PCR fragments were mixed and a second PCR was performed using PenoA_Fw + PenoA_Rv primers. The amplified DNA fragment was digested with *Pst*I and *Xho*I and inserted into *Pst*I/*Sal*I-digested pNGAG1, resulting in pNPenoAΔ−121 to −224GUS.

The pNPenoAΔ−181 to −195GUS plasmid was constructed as follows: PenoAΔ−181 to −195 fused to the *uidA* CDS region was amplified by fusion PCR using pNPenoAGUS and PenoA_Fw + PenoAΔ−181 to −195_Rv and PenoAΔ−181 to −195_Fw + uidA_Rv primers. The amplified DNA fragment was digested with *Pst*I and *Xba*I and inserted into *Pst*I/*Xba*I-digested pNGAG1. pNPenoAΔ−137 to −179GUS was constructed using the same method. PenoA_Fw + PenoAΔ−137 to −179_Rv, and PenoAΔ−137 to −179_Fw + uidA_Rv primers were used to amplify the DNA insert.

Plasmid DNAs used to examine the effect of site-specific mutagenesis in CE_1 and CE_2 on the expression level were constructed as follows: each of the mutations except for Mut 2 mutation in CE_1 was introduced into pNPenoAGUS using the PCR mutagenesis method (described below), obtaining pNPenoAK/Mm1GUS, pNPenoAmCS1GUS, pNPenoAmCS2GUS, pNPenoAmCS3GUS, pNPenoAmCS4GUS, and pNPenoAmCS5GUS. Mut 2 mutation in CE_1 was introduced into pNPenoAK/Mm1GUS using PCR mutagenesis method, obtaining pNPenoAK/Mm2GUS.

Plasmid DNAs used to examine the effect of the intron deletion, 3′ splicing site mutation (3′ ssm), and 5′ splicing site mutation (5′ ssm) in 5′ UTR on the expression level were constructed as follows: DNA fragment of the *enoA* promoter, including the deleted 440 bp intron or 3′ ssm, was amplified by PCR using *A. oryzae* RIB40 genomic DNA and PenoA_Fw + PenoAΔi_Rv or PenoA_Fw + PenoA3′ ssm_Rv, respectively. Each amplified DNA fragment was digested with *Pst*I and *Xho*I and inserted into *Pst*I/*Sal*I-digested pNGAG1, obtaining pNPenoAΔiGUS or pNPenoA3′ssmGUS. 5′ ssm was introduced into pNPenoAGUS using the PCR mutagenesis method, obtaining pNPenoA5′ssmGUS.

Plasmid DNAs used to examine 5′ UTR replacement effects on expression level were constructed using PenoA between −1 and −1,000 fused to the *uidA* CDS region. The insert was amplified by fusion PCR using pNPenoAGUS and PenoAWT_5UTR_Fw + PenoAWT_5UTR_Rv, and *uidA*_Fw + *uidA*_Rv primers. The amplified DNA fragment was digested with *Pst*I and *Xba*I and inserted into *Pst*I/*Xba*I-digested pNGAG1, yielding pNPenoAWT_5UTRGUS. The 5′ UTR replaced PenoA (PenoArDown-Up_5UTR or PenoArUp-Down_5UTR) fused to the *uidA* CDS region were amplified by fusion PCR using pNPenoAWT_5UTRGUS template. The primer sets PenoA if_Fw + PenoArDown-Up_5UTR_Rv and PenoArDown-Up_5UTR_Fw + *uidA*_Rv were used to amplify the PenoArDown-Up_5UTR-*uidA* fragment. PenoA-if_Fw + PenoArUp-Down_5UTR_Rv and PenoArUp-Down_5UTR_Fw + *uidA*_Rv primers were used to amplify PenoArUp-Down_5UTR-*uidA* one. A host vector fragment amplified using pNPenoAGUS DNA and pNGAG1-if_Fw and pNGAG1-if_Rv primers. An In-Fusion HD Cloning Kit (Bio Inc.) was used to insert the constructed 5′ UTR inserts, yielding pNPenoArDown-Up_5UTRGUS and pNPenoArUp-Down_5UTRGUS.

In the *niaD*$^{-}$ strains, the 3′ region of *niaD* CDS was deleted (unpublished). The pUC-niaD plasmid (unpublished) was used to complement the *niaD* mutation by homologous recombination. This plasmid was constructed by inserting the 3′ half of the *niaD* locus region at +1609 to +3876 into *Sma*I-digested pUC119 (Takara Bio Inc.).

Introducing site-specific mutagenesis into plasmid DNA

PCR mutagenesis was used to introduce site-specific base substitution mutagenesis into plasmid DNA. Two complementary primers containing mutated sites flanked by 15−25 bp were used to amplify template plasmid DNA. Primers used for PCR mutagenesis are listed in Supplementary Table 1. The PCR products were digested with *Dpn*□ (NEB) to selectively cut only the template plasmid DNA. The remaining nascent plasmids were then incorporated into *E. coli*. The plasmids were sequenced to verify they contained the desired mutations. Q5 High-Fidelity DNA polymerase (NEB) was used for PCR mutagenesis. All mutated plasmids, except for pNPenoAK/Mm2GUS, were generated from pNPenoAGUS. pNPenoAK/Mm2GUS was generated from pNPenoAK/Mm1GUS.

Construction of DNA fragment for gene disruption

A DNA fragment for *A. oryzae acuK* or *acuM* ortholog gene disruption was constructed by fusion PCR using the *A. nidulans* ATP sulfurylase gene (*sC*) as a selectable marker. DNA containing the *sC* expression construct was amplified by PCR using a pUSC plasmid (Yamada et al., 1997) and AnsC_Fw and AnsC_Rv primers. DNA fragments upstream of the *acuK* CDS and the inner CDS region were amplified using *A. oryzae* genomic DNA and up-acuK_Fw + up-acuK_Rv and CDS-acuK_Fw + CDS-

acuK_Rv primers. The three amplified fragments were mixed and a second PCR was performed using up-acuK_Fw + CDS-acuK_Rv primers, yielding an *acuK* disruption fragment. DNA fragments up- and downstream of the *acuM* CDS were amplified using *A. oryzae* genomic DNA and up-acuM_Fw + up-acuM_Rv and down-acuM_Fw + down-acuM_Rv primers. The two fragments and the *sC*-fragment were mixed and a second PCR was performed using up-acuM_Fw + down-acuM_Rv primers, resulting in an *acuM* disruption fragment.

Construction of DNA fragment for *enoA* promoter replacement

A DNA fragment for *A. oryzae enoA* promoter replacement was constructed by multiple fragment cloning of PCR products using the In-Fusion HD Cloning kit (Takara Bio USA Inc.). DNA fragment containing the *A. nidulans sC* gene as a selectable marker was amplified by PCR using the plasmid pUSC (Yamada et al., 1997), and AnsC_Fw and AnsC_Rv primers. DNA fragments of the proximal and distal 5′-flanking regions of the *enoA* gene were amplified using *A. oryzae* RIB40 genomic DNA with Up-PenoA-if_Fw + Up-PenoA-if_Rv and PenoA_if_Fw2 + PenoA_if_Rv, respectively. The distal *enoA* 5′-flanking region contained the 3′-flanking region of the adjacent gene *AO090003000054* (see Fig. 4). The three amplified fragments were cloned into pUC19, resulting in pCPeR. 5′ ssm was introduced into the *enoA* promoter region in pCPeR by the PCR mutagenesis method, resulting in pCPe5ssmR. DNA fragment was amplified by PCR using pCPeR or pCPe5ssmR with Up-PenoA-if_Fw + PenoA_if_Rv2, and then used for replacement of the native *enoA* promoter. Similarly, replacement by the *enoA* promoter with mCS3 mutation was performed using the DNA fragments amplified by PCR using pNPenoAmCS3GUS and *A. oryzae* RIB40 genomic DNA as templates with

315 PenoA_if_Fw2 + PenoA_if_Rv2 and CDS-enoA-if_Fw + CDS-enoA-if_Rv primers,
316 respectively, resulting in pCPemCS3R.
317
318 **Transformation experiments**
319 *E. coli* and *A. oryzae* were transformed as previously described (Inoue et al., 1990,
320 Gomi et al., 1987).
321
322 **Southern blot analysis**
323 *A. oryzae* plasmid insertion, gene disruption, and promoter replacement were
324 confirmed by southern blot (data not shown). Genomic DNA preparation and southern
325 blot analysis were performed using the method described by Tanaka et al. (2012).
326 Transformant DNA containing *uidA* expression constructs were digested with *Pst*I. A
327 probe was amplified with niaD-probe_Fw + niaD-probe_Rv primers. To analyze *acuK* or
328 *acuM* disruptants, each genomic DNA was digested with *Xba*□ or *Pst*□ and each probe
329 was amplified using acuK-probe_Fw + acuK-probe_Rv or acuM-probe_Fw + acuM-
330 probe_Rv primers. To confirm *enoA* promoter replacement, genomic DNA was digested
331 with *Eco*R□ and a probe was amplified with enoA-CDS_Fw + CDS-enoA-if_Rv primers.
332
333 **β-glucuronidase (GUS) reporter assay**
334 Mycelia of *A. oryzae* transformants containing *uidA* expression constructs were
335 disrupted in liquid nitrogen using a mortar and pestle. Mycelial extracts were prepared
336 using the method described by Tada et al. (1991). Protein concentration of the mycelial
337 extracts was measured by Bradford assay (1976) using bovine serum albumin as a
338 standard. Quantitative GUS activity analysis was performed by spectrophotometry using

the modified method of Jefferson et al. (1989). Samples were mixed with 800 µl buffer (0.2% Triton X-100, 100 mM NaH_2PO_4, pH 7.0) containing 10 mM *p*-nitrophenyl-β-D-glucuronide substrate and incubated at 37□ for 20 min. The reaction was terminated by adding 320 µl 1 N sodium hydroxide. The *p*-nitrophenol absorbance was measured at 415 nm. One unit was defined as the amount of enzyme required to produce one nanomole *p*-nitrophenol per min.

Computational MEME analysis for consensus motif discovery

The 5′-flanking regions 1,000 bp from enolase encoding regions in four *Aspergillus* species, *A. oryzae*, *A. nidulans*, *Aspergillus niger*, and *Aspergillus fumigatus* were collected from the *Aspergillus* Genome Database (http://www.aspgd.org/) and used as an enolase gene promoter data set. Consensus motifs were queried using the MEME algorithm (Bailey et al., 2009), using the data set and the MEME Suite Software Web server (Bailey et al., 2009; http://meme-suite.org/tools/meme). Motifs with *E*-values < 0.05 were considered statistically significant consensus motifs.

Results:

Identification of two transcription start sites (TSSs) in *enoA*

While aligning EST data (Akao et al., 2007, https://nribf21.nrib.go.jp/EST2/) and *A. oryzae* genome sequencing data (Machida et al., 2005), we recognized the possibility of two 5′ ends in *enoA* transcripts. To examine whether *enoA* indeed has multiple TSSs, we used 5′ SAGE to identify putative TSSs, which indicated the presence of two TSSs (Fig. 1A). The upstream TSS (uTSS) was located around −510 relative to the start codon (+1), while the downstream TSS (dTSS) was located around −35 (Fig 1A, B). EST sequence

363 analysis also revealed a 440 bp intron within the 5′ UTR when transcription was initiated
364 at uTSS (Fig. 1B). Interestingly, the EST occurrence pattern of the two TSSs differed
365 under two culture conditions–liquid nutrient-rich culture (LR) and solid-state culture with
366 wheat bran (SW) (Table 1, Akao et al., 2007). ESTs derived from the dTSS were mainly
367 obtained in LR cultures, while ESTs derived from uTSS were obtained exclusively from
368 SW cultures (Table 1). These data suggest that the usage of two *enoA* TSSs is altered by
369 varying culture conditions. The EST data also suggest that another glycolytic pathway
370 gene, *gpdA*, which encodes glyceraldehyde-3-phosphate dehydrogenase (GAPDH),
371 seems to use alternative TSSs depending on culture conditions (Table 1).
372
373 **_enoA_ alternative TSS selection depends on carbon source types associated with**
374 **glycolysis and gluconeogenesis**
375 Because uTSS was exclusively selected under solid-state culture conditions with
376 limited carbon sources available for *A. oryzae*, we hypothesized that distinct carbon
377 sources may affect which TSS drives transcription of the *enoA* gene. Therefore, we
378 examined whether alternative TSS selection depends on the available carbon source
379 species. We used qRT-PCR analysis using primers designed to discriminate the *enoA*
380 transcript variants derived from the two TSSs (Fig. 1B). An *A. oryzae* wild-type strain,
381 RIB40, was grown in liquid medium containing multiple carbon sources (glucose,
382 fructose, glycerol, acetate, and ethanol). The ratio of TSS usage in *enoA* was then
383 calculated (Fig. 1C). The uTSS and dTSS transcript levels were 0.03–0.2 and 0.7–0.9
384 relative to the total *enoA* transcript, respectively, when grown with carbon sources that
385 are metabolized through glycolysis (Fig. 1C). In contrast, predominant uTSS usage was
386 evident when grown in acetate and ethanol, which are potential substrates for

gluconeogenesis (Fig. 1C). These results suggest that the dTSS is predominantly used under glycolytic conditions, while the uTSS is preferentially used under gluconeogenic conditions. Moreover, when grown in glycerol, which is metabolized in both glycolysis and gluconeogenesis, dTSS and uTSS were not preferentially used.

RT-PCR analysis showed that highly efficient intron splicing within the 5′ UTR occurred in the *enoA* primary transcripts derived from the uTSS (Fig. 1D). The dTSS-derived transcript was not affected by an unspliced transcript derived from the uTSS. Furthermore, northern blot analysis showed that the total *enoA* transcript level varies depending on the carbon source (Fig. 1E). These results suggest that the selection between the two *enoA* TSSs is dependent on metabolic state, i.e. glycolysis or gluconeogenesis. Further, alternative TSSs usage is associated with transcript level control.

Alternative TSSs usage does not affect *enoA* translational efficiency or primary protein structure

Next, we clarified the functional significance of alternative TSSs selection. It is possible that the use of several TSSs can generate diverse protein structures in eukaryotes (Ayoubi and Van De Ven, 1996). Although there was an intron within the *enoA* 5′ UTR, no upstream start codons were observed in either 5′ UTR derived from the two TSSs. Therefore, the *enoA*-encoded primary protein structure is not affected by the presence of two TSSs. However, 5′ UTR heterogeneity could affect translation efficiency (Davuluri et al., 2008; Rojas-Duran and Gilbert, 2010). The use of two TSSs in *enoA* produces two distinct 5′ UTRs (Fig. 1B) that lack upstream ORFs that could have serious adverse effects on translation (Morris and Geballe, 2000), though both the length and sequence differ. However, GUS reporter analysis of the *enoA* promoter plus 5′ UTR revealed that

411 replacing each 5′ UTR with another 5′ UTR did not alter *enoA* promoter activity (Fig. S1).
412 This suggests that alteration of the 5′ UTR caused by differential TSS use does not affect
413 *enoA* translation efficiency.
414
415 **AcuK and AcuM transcription factors upregulate *enoA* transcription from the uTSS**
416 Because the *enoA* gene uses carbon source-dependent alternative TSSs for
417 transcription, we investigated the molecular mechanism of *enoA* transcriptional
418 regulation. To identify putative *cis*-elements for *enoA* transcription from the two distinct
419 TSSs, we analyzed the 5′-flanking region 1,000 bp upstream of the enolase-encoding gene
420 start codon in four *Aspergillus* species (*A. oryzae*, *A. nidulans*, *A. fumigatus*, and *A. niger*)
421 by *in silico* motif prediction using MEME Suite software (Bailey et al., 1999). Two
422 consensus sequences, designated CE_1 and CE_2, were identified in all *Aspergillus*
423 promoters with *E*-values < 1×10^{-10} (Fig. 2A, Fig. S2A). The CE_1 sequence was located
424 −761 and −724 upstream of the uTSS in the *enoA* promoter of *A. oryzae*. Interestingly,
425 the putative binding motif (CG(C)GN$_7$CG(C)G) of the key transcription factors AcuK
426 and AcuM were identified within this region (Fig. 2A, B). These transcription factors
427 were identified in *A. nidulans* and regulate gluconeogenesis (Hynes et al., 2007: Suzuki
428 et al., 2012). Because uTSS selection seems to depend on the gluconeogenic carbon
429 sources, AcuK and AcuM could be involved in transcription from the uTSS. Disrupting
430 the *acuK* or *acuM* orthologous gene in *A. oryzae* resulted in reduced growth on acetate
431 but not on glucose (Fig. S2B). In both Δ*acuK* and Δ*acuM* disruptants, transcript levels of
432 genes involved in gluconeogenesis, such as phosphoenolcarboxykinase-encoding gene
433 (*pckA*) and fructose-1,6-bisphosphatase-encoding gene (*fbpA*), were significantly
434 decreased when cultured in acetate-containing media (Fig. S2C). The Δ*acuK* and Δ*acuM*

phenotypes in *A. oryzae* are consistent with similar *A. nidulans* phenotypes (Suzuki et al., 2012). qRT-PCR analysis of Δ*acuK* and Δ*acuM* showed significantly reduced *enoA* transcripts derived from the uTSS in acetate-culture conditions (Fig. 2D). In addition, site-specific mutations in the AcuK and AcuM binding motifs resulted in a significant decrease in *enoA* promoter activity (Fig. 2F). Conversely, culture media containing glucose did not reduce AcuK and AcuM motif-related *enoA* promoter activity (Fig. 2E). In the transformant expressing *uidA* by the *enoA* promoter with Mut 2 mutation (Fig. 2C), uTSS-derived *uidA* transcript level was significantly decreased in acetate-culture condition (Fig. S3A). These results indicate that *acuK* and *acuM* are required for gluconeogenesis and are involved in *enoA* transcription from the uTSS in gluconeogenic conditions in *A. oryzae*. However, the deletion of *acuK* or *acuM* did not completely abolish *enoA* gene expression (Fig. 2D), suggesting that other regulators are involved in the transcription at uTSS.

Identification of *cis*-elements required for *enoA* transcription from the dTSS in glycolytic conditions

In addition to the conserved sequence CE_1, which encompasses putative AcuK and AcuM binding motifs, another conserved sequence, CE_2, was identified by *in silico* motif screening using enolase promoters from four *Aspergilli*. The CE_2 sequence is located −178 to −154 upstream of the start codon in the *A. oryzae enoA* promoter (Fig. 2B). Notably, CE_2 was detected within the region between −224 and −121. Deletion of this region results in drastically decreased promoter activity under glucose culture conditions (Toda et al., 2001). The 15 bp sequence located at −195 to −181 is a *cis*-regulatory element according to EMSA analysis using whole-cell extracts (Toda et al.,

2001). However, the importance of CE_2 has not been reported. We performed additional *enoA* promoter deletion analysis to determine which element is involved in *enoA* expression. Deletion of the 15 bp region located at −195 to −181 resulted in ~35% decrease in promoter activity, whereas deletion of the 42 bp region located at −179 to −137 reduced promoter activity by ~85% under glucose culture conditions. Similarly, deleting the 104 bp region from −224 to −121 was nearly equivalent (~90%) (Fig. S4). Furthermore, site-specific mutations at four independent consensus sites (mCS1 to mCS4) in the CE_2 sequence (Fig. 2G) resulted in a significant decrease in promoter activity under glucose culture conditions. The mCS5 mutation outside of the CE_2 sequence caused no substantial change in the promoter activity (Fig. 2H). The mCS3 promoter mutation had the lowest GUS activity, slightly lower than deletion of the 42 bp region located at −179 to −137 (Fig. 2H). In addition, mCS3 mutation resulted in a significant decrease in dTSS-derived *uidA* transcript level in glucose-culture condition (Fig. S3B). These data suggest that the CGG sequence is required for *enoA* transcription. However, the mCS3-containing promoter showed no significant decrease in promoter activity in acetate culture conditions (Figs. 2I and S3B). These results indicate that a crucial *cis*-regulatory element involved in transcription from the dTSS under glycolytic conditions is contained within the CE_2 sequence, but does not include the previously-described 15 bp sequence (Toda et al., 2001). Furthermore, introducing the mutation mCS3 into the endogenous *enoA* promoter resulted in reduced growth on glucose-containing agar medium but not on acetate-containing agar medium (data not shown). Therefore, our promoter analyses demonstrate that two conserved sequences among *Aspergilli*, CE_1 and CE_2, function as *cis*-elements in *enoA* transcription from the uTSS under gluconeogenic conditions and from the dTSS under glycolytic conditions.

Effects of the 5′ UTR intron on *enoA* gene expression

In *enoA*, the 440 bp sequence is spliced as an intron within 5′ UTR on transcription initiation at uTSS (Fig. 1B, D). The intron length seems noticeably long in filamentous fungi including *Aspergilli*. To investigate the significance of the intron within 5′ UTR, we examined the effect of the mutation in the 5′ or 3′ splice site and intron deletion on the *uidA* reporter activities (Fig. 3A). GUS activity was unaffected by 5′ ssm and 3′ ssm under glucose-culture condition, while it decreased drastically under acetate-culture condition (Figs. 3B and 3C). Northern blot analysis showed that longer *uidA* transcripts presumably derived from unspliced mRNAs were detected in 5′ ssm and 3′ ssm under acetate-culture condition (Fig. 3D). The RT-PCR experiment confirmed the presence of an unspliced mRNA in 5′ ssm and 3′ ssm (data not shown). This suggests that splicing of the intron within 5′ UTR is essential for efficient translation from the uTSS-derived transcript. As expected, deletion of the intron led to loss of both GUS activity and *uidA* transcript under glucose-culture condition (Figs. 3B and 3D) because of the elimination of dTSS within the intron. However, the intron deletion resulted in a significant decrease in GUS activity as well as *uidA* transcript level under acetate-culture condition (Figs. 3B and D), suggesting that the intron within 5′ UTR increases uTSS-derived transcript level. Furthermore, combination of the intron deletion and mutation in AcuK/AcuM binding motif resulted in a substantial loss of GUS activity in acetate-culture condition (Fig. 3E). This suggests that the intron and AcuK/AcuM independently enhance the gene expression from uTSS.

Physiological significance of alternative TSS usage in *enoA*

To examine the physiological significance of alternative TSS usage in *enoA*, we generated a transformant in which the native *enoA* promoter was replaced with the promoter containing mCS3 or 5′ ssm (Fig. 4A). The transformant harboring mCS3 in the *enoA* promoter (mCS3 strain) showed a significantly poor growth in glucose-culture condition, and the transformant harboring 5′ ssm (5′ ssm strain) could hardly grow in acetate-culture condition (Fig. 4B). These results strongly support that transcriptional induction from dTSS and uTSS in *enoA* are crucial for *A. oryzae* growth on glucose and acetate, respectively.

Prevalence of alternative TSS usage in glycolytic/gluconeogenic genes

We next examined the presence of alternative TSSs in other glycolytic pathway genes, including *gpdA*, which has alternative TSSs (Table 1). EST and 5′ SAGE analysis indicated two TSSs located at −164 or −74 and a 104-bp intron is present within the *gpdA* 5′ UTR (Table 1, Fig. 5A). Contrary to findings in *enoA*, EST data showed almost the same occurrence of two TSSs in LR conditions, whereas dTSS-mediated transcription was significantly decreased in the SW condition (Table 1). The ratio of dTSS and uTSS transcripts relative to the total *gpdA* transcripts showed that the uTSS is predominantly used in acetate- and glucose-culture conditions (Fig. 5B), suggesting that *gpdA* TSS selection is less stringent than in *enoA*. 5′ SAGE analysis indicated a single TSS in the *gpdB* gene, a *gpdA* paralog (Fig. S5A).

Because there was insufficient data to identify TSSs obtained from EST and 5′ SAGE analyses, we performed 5′ RACE analysis in other 7 glycolytic pathway genes in glucose and acetate culture conditions. Our analysis suggested the presence of multiple TSSs in 5 glycolytic pathway genes. Particularly, the *fbaA* gene encoding fructose-bisphosphate

aldolase showed stringent selection of two TSSs in response to carbon sources similar to *enoA*. *fbaA* also contained a 270 bp intron within its 5′ UTR (Fig. 5C). qRT-PCR analysis demonstrated that *fbaA* transcription starts exclusively from the uTSS located around −380 under gluconeogenic conditions and from the dTSS located around −70 under glycolytic conditions. Possible alternative TSSs were present in other 3 genes, including *pgkA* (phosphoglycerate kinase), *gpmA* (phosphoglycerate mutase), and *tpiA* (triose-phosphate isomerase), whose transcription seemed to start from the uTSS in the presence of acetate but not glucose. In contrast, the *pgiA* gene (glucose-6-phosphate isomerase) also appeared to have alternative TSSs, but TSS selection did not depend on carbon source. Interestingly, all glycolytic pathway genes with putative alternative TSSs, except for the *tpiA* gene, contained an intron within their 5′ UTRs (Fig. S5).

Among the glycolytic pathway genes tested, only one gene, *pkiA* (pyruvate kinase), did not have alternative TSSs (Fig. S5). Most glycolytic enzymes catalyze reversible reactions in glycolysis and gluconeogenesis. Indeed, almost all the genes we tested that encode enzymes catalyzing reversible reactions use alternative TSSs, except for *gpdB*. In this context, it is interesting that *pkiA*, which encodes an enzyme catalyzing an irreversible glycolytic reaction, has a single TSS for its transcription. Therefore, we examined the TSSs of other genes involved in irreversible glycolysis and gluconeogenesis reactions. *pfkA* encodes phosphofructokinase, and was expressed only under glycolytic conditions. Fructose-1,6-bisphosphatase-encoding *fbpA* and phosphoenolpyruvate carboxykinase-encoding *pckA* were expressed only under gluconeogenic conditions (data not shown). 5′ RACE analysis was performed on *pfkA*, *fbpA*, and *pckA* in the presence of glucose or acetate alone. All the three genes had a single TSS. Pyruvate carboxylase-endcoding *pycA* plays an important role in gluconeogenesis, and was expressed in both glucose- and

555 acetate-culture conditions. *pkiA* was expressed in a similar manner (data not shown).
556 Interestingly, 5′ RACE analysis showed that *pycA* had multiple TSSs and used uTSS and
557 dTSS under glucose- and acetate-culture conditions, respectively. This trend was opposite
558 from alternative TSS usage in other tested genes (Fig. 5D). Furthermore, alternative
559 splicing occurred within the 5′ UTR in *pycA* primary transcripts derived from the uTSS,
560 resulting in three alternatively-spliced transcript variants (Fig. 5D).

561 Alternative TSSs were observed in most glycolytic pathway genes involved in
562 reversible reactions. These genes also contained an intron within the 5′ UTR in uTSS-
563 derived primary transcripts in *A. oryzae* (Fig. 5E). This suggests that the use of alternative
564 TSSs is not unique to *enoA*. Rather, alternative TSSs are common among glycolysis and
565 gluconeogenesis genes to some extent, although the alternative TSS usage pattern in *enoA*
566 and *fbaA* appears to depend on glycolytic or gluconeogenic carbon sources.

567

568 **Usage of two TSSs in enolase-encoding genes differs between *A. oryzae* and *A.**
569 ***nidulans* under glycolytic conditions**

570 In *A. nidulans*, it is possible that the enolase-encoding gene *acuN* also has two TSSs–
571 a uTSS located at −426 and a dTSS located at −4 (Hynes et al., 2007). Furthermore, the
572 *acuN356* mutation, with a break point at −220, results in loss of growth on gluconeogenic
573 carbon sources but not on glycolytic carbon sources (Armitt et al., 1976; Hynes et al.,
574 2007). These observations suggest that the *A. nidulans acuN* gene is also transcribed
575 preferentially from the uTSS under gluconeogenic conditions and from the dTSS under
576 glycogenic conditions, similar to *A. oryzae enoA*. To address this possibility, we first
577 confirmed TSSs in *A. nidulans acuN* by 5′ RACE analysis. The *acuN* gene possessed two
578 TSSs located around −440 and −20, consistent with previous studies (Hynes et al., 2007)

579 (Fig. 6A). In addition, 5′ end clones obtained from acetate-culture conditions were
580 transcribed from the uTSS. Further, an intron of 385 bp is present within the 5′ UTR of
581 the primary transcript (Fig. 6A). Unexpectedly, in glucose-culture conditions, 5′ end
582 clones derived from the dTSS were not predominant (Fig. 6A). qRT-PCR analysis was
583 performed to estimate the TSSs usage ratio, which showed that the uTSS- and dTSS-
584 derived transcripts relative to the total *acuN* transcripts were 0.4−0.6 and 0.2−0.3,
585 respectively, in the presence of glycolytic carbon sources. In contrast, *acuN* transcription
586 occurred exclusively from the uTSS under gluconeogenic conditions (Fig. 6B).
587 Furthermore, northern blot analysis showed that *acuN* was transcribed at higher level in
588 the presence of acetate and ethanol than in the presence of glucose and fructose (Fig. 6C).
589 These data indicate that total *acuN* transcript levels in glycolytic and gluconeogenic
590 carbon sources could be different from *enoA*, which was more highly expressed in the
591 presence of glucose and fructose (Fig. 1E, Fig. 6D). Clearly, TSSs usage in enolase-
592 encoding genes is divergent between *A. oryzae* and *A. nidulans* (Fig. 6D), although highly
593 conserved *cis*-element sequences required for gene expression exist upstream of the uTSS
594 and dTSS in *enoA* and *acuN*.
595
596 **Discussion:**
597 Glycolysis is a fundamental metabolic pathway for cellular energy acquisition. In *A.*
598 *oryzae*, an industrially important filamentous fungus, glycolytic genes are strongly
599 expressed at the transcriptional level in the presence of fermentable carbon sources.
600 Although this transcriptional profile may be important for growth in fermentative culture
601 conditions, the details of transcriptional regulation in glycolytic genes remain to be
602 elucidated.

We investigated molecular transcriptional control in the enolase-encoding gene *enoA*, which is strongly expressed in *A. oryzae*, focusing on TSS regulation. We demonstrated the presence of two TSSs in *enoA* and that TSS selection appears to be strictly dependent on the carbon source metabolized via glycolysis or gluconeogenesis. Furthermore, *enoA* transcript levels depend on the carbon source. *enoA* is more highly expressed with glycolytic carbon sources than gluconeogenic carbon sources (Fig. 1E). Because neither the *enoA* protein primary structure nor translation efficiency was affected by alternative TSS usage, it is possible that *enoA* alternative TSSs play an important role in transcriptional regulation in response to available environmental carbon sources. Thus, to elucidate the molecular details of *enoA* transcriptional regulation using alternative TSSs, we identified *cis*-regulatory elements in the *enoA* promoter and found that highly conserved sequences are present in enolase-encoding gene promoters among *Aspergilli* (Fig. 2). CE_1 encompasses the AcuK and AcuM binding motif responsible for gluconeogenic gene expression. Mutations in this motif result in a significant decrease in *enoA* promoter activity, indicating its importance for uTSS-initiated *enoA* transcription in gluconeogenic conditions. The function of the second highly-conserved sequence contained in CE_2 remains unclear, but mutation analyses showed that this sequence is involved in dTSS-initiated *enoA* transcription in glycolytic conditions. It has not yet been determined whether the CE_2 sequence is also required for enolase-encoding gene expression in other *Aspergilli*. Additionally, the sequence of the *cis*-regulatory element in CE_2 remains to be identified. Further studies are required to understand the significance of the CE_2 sequence, identify the regulatory *cis*-element in CE_2 by EMSA, and define which regulatory protein(s) binding to this sequence.

Enolase catalyzes the reversible conversion of 2-phosphoglycerate to

phosphoenolpyruvate in glycolysis and gluconeogenesis. Transcription of the *A. oryzae enoA* gene can be initiated from different TSSs depending on glycolytic or gluconeogenic carbon sources, suggesting that alternative TSS use is a characteristic feature of glycolytic pathway genes. Although all the glycolytic/gluconeogenic genes could not be investigated, most genes involved in reversible reactions likely have multiple TSSs and contain an intron within uTSS-derived primary transcripts. However, carbon source-dependent alternative TSS use is not conserved, except in *enoA* and *fbaA*. Thus, although most glycolytic/gluconeogenic genes contain alternative TSSs, their usage patterns are not regulative. Nevertheless, *fbaA* showed predominant uTSS and dTSS use under gluconeogenic and glycolytic conditions, respectively, similar to *enoA*. This characteristic feature is supported by qRT-PCR analysis showing that the *fbaA* transcripts derived from the uTSS and dTSS were exclusively obtained in the presence of acetate and glucose, respectively (data not shown). Like *enoA*, *fbaA* also contains a relatively long intron (229 bp) within its uTSS-derived primary transcript. Furthermore, a putative AcuK and AcuM binding motif ($CGGN_7CGG$) was present upstream of uTSS in the *fbaA* promoter region. Mutations in the binding motif significantly decreased *fbaA* promoter activity in the presence of acetate but not glucose (data not shown). Similarly, sequences highly homologous to the conserved CGGTGAA sequence were present upstream of dTSS in the *fbaA* promoter. Further, mutations in these sequences resulted in a significantly decreased *fbaA* promoter activity in the presence of glucose but not acetate (data not shown). These results suggest that the AcuK/AcuM binding motif and *enoA* CE_2 consensus sequences are involved in *fbaA* transcription from the uTSS and dTSS under gluconeogenic and glycolytic conditions, respectively. However, the AcuK/AcuM binding motif was also present upstream of uTSSs and a CGGTGAA-like sequence was

found upstream of dTSSs in most glycolytic genes, suggesting that these element sequences are required for glycolytic gene expression, but not enough to stringently regulate carbon source-dependent uTSS or dTSS selection. It would be an interesting challenge to identify putative *cis*-elements or transcriptional regulators involved in stringent alternative TSS selection by glycolytic or gluconeogenic conditions.

Introduction of the splicing site mutations in *enoA* 5′ UTR resulted in a drastic reduction in GUS activity despite the presence of transcripts (Figs. 3C and 3D). This suggests that intron splicing within 5′ UTR is essential for efficient translation from the uTSS-derived transcript in *enoA*. Three upstream ORFs (uORFs) can be found within the unspliced 5′ UTR sequence and these uORFs might interfere with the translation from the transcript. Moreover, the *fbaA*, *pgiA,* and *acuN* genes contain such cryptic uORFs in their 5′ UTR intron sequences. Although the significance of introns in these genes is unclear, intron splicing may be important for preventing the emergence of uORFs in 5′ UTR. However, deletion of the intron within 5′ UTR resulted in a significant decrease in gene expression from uTSS (Fig. 3C), suggesting that the intron contributes to an increase in uTSS-derived transcript level. Intron-dependent enhancement (IDE) in gene expression has been shown in several eukaryotic organisms, but the molecular mechanisms seem to be divergent across genes or species (Agarwal and Ansari, 2016; Bicknell et al., 2012; Goebels et al., 2013; Rose et al., 2011). Elucidation of the specific molecular mechanisms of IDE in *enoA* would be an important challenge to understand the molecular mechanisms of IDE in *Aspergilli*.

Although the *A. nidulans* enolase-encoding gene *acuN* also has alternative TSSs and a long intron within uTSS-derived primary transcripts similar to *A. oryzae enoA*, alternative TSS selection in *acuN* appears to be less dependent on glycolytic and

675 gluconeogenic carbon sources. *A. oryzae fbaA* showed stringent alternative TSS selection
676 depending on the available carbon source, similar to *enoA*. While we did not investigate
677 alternative TSSs in other *A. nidulans* glycolytic genes, *A. nidulans fbaA* likely also has
678 alternative TSSs. This hypothesis is supported by the observation that the *fbaA1013* strain
679 contains a translocation mutation in an intron within the 5′ UTR (Roumelioti et al., 2010).
680 Thus, future studies examine *fbaA* transcription in *A. oryzae* and *A. nidulans* to compare
681 the regulatory mechanisms in glycolytic genes.

682 Furthermore, despite the presence of highly conserved CE_1 and CE_2 sequences in
683 both *enoA* and *acuN* promoters, *acuN* was highly expressed in the presence of
684 gluconeogenic carbon sources, whereas *enoA* expression occurred in the presence of
685 glucose. These differences in *enoA* and *acuN* transcription might reflect phylogenetic
686 diversity between *A. oryzae* and *A. nidulans*. *A. oryzae* grows rapidly in fermentable
687 carbon sources such as glucose, with much higher maximum specific growth rate in
688 glucose-containing batch cultivations than *A. niger* and *A. nidulans* (Anderson et al.,
689 2008). In general, glycolysis is a critical first step in energy production in living organisms.
690 Thus, higher dTSS-induced *enoA* gene expression in the presence of glucose is associated
691 with *A. oryzae*, which can grow quickly in glycolytic conditions.

692 *A. oryzae* was domesticated from an atoxigenic strain of the ancestor species
693 *Aspergillus flavus* by artificial selection of industrially suitable fungal strains for
694 traditional Japanese fermented food production (Gibbons et al., 2012; Gibbons and Rokas,
695 2013). In sake production, *A. oryzae* is grown on steamed rice grain, producing large
696 amounts of amylolytic enzymes, which degrade rice starch to glucose (Machida et al.,
697 2008; Gomi, 2019). The intrinsic capability of *A. oryzae* to degrade rice starch correlates
698 with 2 or 3 copies of the α-amylase (TAA) gene that was highly expressed among *A.*

699 *oryzae* genes (Hunter et al., 2011; Gibbons et al., 2012). In contrast, a single TAA gene is
700 present with lower expression in the *A. flavus* ancestor (Gibbons et al., 2012). These facts
701 suggest that during domestication, *A. oryzae* was adapted to efficiently assimilate glucose
702 in growth environments on steamed rice. This may explain the high *enoA* transcript level
703 in the presence of glucose. In addition, the stringent selection of alternative TSSs in *enoA*
704 and *fbaA* may be associated with the adaptation to starch-rich growth conditions, although
705 the evolutionary advantages of stringent alternative TSS selection are unclear. Based on
706 the significantly high similarity (99.5%) between the *A. oryzae* and *A. flavus* genomes
707 (Payne et al., 2006; Gibbons et al., 2012), *A. flavus enoA* promoter sequence would be
708 very similar *A. oryzae*. To assess the hypothesis that the domestication process may alter
709 glycolytic gene transcriptional patterns, at least of the *enoA* gene, it would be interesting
710 to examine the transcriptional features of *enoA* in *A. flavus*. Additionally, TAA transcript
711 levels were higher in the *A. oryzae* RIB40 strain used here than in any other *A. oryzae*
712 strains examined (Gibbons et al., 2012). Therefore, high *enoA* expression in the presence
713 of glucose may be specific to *A. oryzae* RIB40. Hence, it is necessary to examine *enoA*
714 expression profiles in other *A. oryzae* strains and in *Aspergillus sojae*, an important koji
715 mold closely related to *A. oryzae* (Sato et al., 2011).
716 Furthermore, it would be interesting to elucidate how the transcriptional pattern of
717 enolase-encoding genes alters between *A. oryzae* and *A. nidulans*. To investigate the effect
718 of the genetic background on enolase gene transcription in the two species, we replaced
719 the endogenous promoter *enoA* with *acuN* in *A. oryzae*. Further, no significant change
720 was observed in TSS usage and transcript level in both glucose- and acetate-culture
721 conditions (data not shown), suggesting that alternative TSS usage patterns between *A.*
722 *oryzae* and *A. nidulans* are dependent on the difference in genetic backgrounds other than

promoter sequences. Further studies are required to identify transcription factors that bind to *cis*-elements and elucidate the manner in which such transcription factors are involved in alternative TSS selection between the two fungal species.

The biological significance of alternative TSSs is revealed in the present study. Indeed, most glycolytic genes possess alternative TSSs. Transcriptional control based on alternative TSSs is not rare in eukaryotic microbes. Comprehensive TSS analyses suggest multiple TSSs in genes in some fungal species such as *S. cerevisiae* (Miura et al., 2006), *Shizosaccharomyces pombe* (Li et al., 2015), *A. nidulans* (Sibthorp et al., 2013), and *Coprinopsis cinerea* (Cheng et al., 2013). In addition, alternative TSS usage occurs in some genes in response to changing physiological conditions, e.g. conidiophore development in *A. nidulans* (Prade and Timberlake, 1993), hyphal growth during sexual development in *Cryptococcus neoformans* (Kaur and Panepinto, 2016), insect infection in *Metarhizium robertsii* (Guo et al., 2017), and zinc homeostasis and meiosis in *S. cerevisiae* (Taggart et al., 2017; Tresenrider and Ünal, 2018). However, reports describing genes with alternative TSSs in fungi are considerably fewer than in mammals (Davuluri et al., 2008; Forrest et al., 2015), because sufficient TSSs data in multiple physiological conditions has not been accumulated despite high environmental adaptability of fungi. More comprehensive analyses on the relationship between gene function and transcriptional patterns are required to better understand the biological significance of alternative TSSs in fungi. We believe that genome-wide comparative analysis of carbon source-dependent TSS usage profiles is the first step to investigate the biological significance of alternative TSS usage in fungi, and particularly in *Aspergillus* spp. We are now planning TSSs analysis in *A. oryzae* and *A. nidulans* using the cap analysis gene expression (CAGE) method (Shiraki et al., 2003).

In conclusion, this study provides evidence that alternative TSS usage in the *A. oryzae* enolase-encoding gene (*enoA*) is stringently observed in glycolytic/gluconeogenic conditions. Moreover, it revealed that two highly conserved sequences in the promoter among *Aspergilli* function as *cis*-regulatory elements for enhancing transcription from two TSSs. Furthermore, the aldolase-encoding gene (*fbaA*) also shows alternative TSS usage similar to *enoA*. These findings can further our understanding about transcriptional regulation of glycolytic/gluconeogenic genes in *A. oryzae*. In addition, our results suggested that alternative TSS usage in enolase-encoding genes could be diversified in *Aspergilli*, despite the presence of well-conserved *cis*-elements. This finding provides novel insights into the diversity of transcriptional regulation of primary metabolic genes in *Aspergilli*. We expect *Aspergillus* to serve as a model group for future studies unraveling the evolutionary significance of alternative TSS usage in fungi.

Conflict of Interest: The authors declare no conflicts of interest.

References:

Agarwal N, Ansari A (2016) Enhancement of transcription by a splicing-competent intron is dependent on promoter directionality. PLoS Genet 12: 20 doi: 10.1371/journal.pgen.1006047

Akao T, Gomi K, Goto K, Okazaki N, Akita O (2002) Subtractive cloning of cDNA from *Aspergillus oryzae* differentially regulated between solid-state culture and liquid (submerged) culture. Curr Genet 41: 275−281 doi: 10.1007/s00294-002-0314-y

Akao T, Sano M, Yamada O, Akeno T, Fujii K, Goto K, Ohasi-Kunihiro S, Takase K, Yasukawa-Watanabe M, Yamaguchi K, et al. (2007) Analysis of expressed sequence tags from the fungus *Aspergillus oryzae* cultured under different conditions. DNA Research 14: 47−57 doi: 10.1093/dnares/dsm008

Andersen MR, Vongsangnak W, Panagiotou G, Salazar MP, Lehmann L, Nielsen J (2008)

775 A trispecies *Aspergillus* microarray: Comparative transcriptomics of three
776 *Aspergillus* species. Proc Natl Acad Sci U S A 105: 4387−4392 doi:
777 10.1073/pnas.0709964105
778 Armitt S, McCullough W, Roberts CF (1976) Analysis of acetate non-utilizing (*acu*)
779 mutants in *Aspergillus nidulans*. J Gen Microbiol 92: 263−282 doi:
780 10.1099/00221287-92-2-263
781 Asai T, Tsukada K, Ise S, Shirata N, Hashimoto M, Fujii I, Gomi K, Nakagawara K,
782 Kodama KN, Oshima Y (2015) Use of a biosynthetic intermediate to explore the
783 chemical diversity of pseudo-natural fungal polyketides. Nat Chem 7:737−743
784 doi: 10.1038/nchem.2308
785 Ayoubi TAY, VanDeVen WJM (1996) Regulation of gene expression by alternative
786 promoters. FASEB J 10: 453−460 doi: 10.1096/fasebj.10.4.8647344
787 Barbesgaard P, Heldt-Hansen HP, Diderichsen B (1992) On the safety of *Aspergillus*
788 *oryzae*: a review. Appl Microbiol Biotechnol 36:569–572.
789 doi:10.1007/bf00183230
790 Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren JY, Li WW, Noble
791 WS (2009) MEME SUITE: tools for motif discovery and searching. Nucleic
792 Acids Res 37: W202−W208 doi: 10.1093/nar/gkp335
793 Bicknell AA, Cenik C, Chua HN, Roth FP, Moore MJ (2012) Introns in UTRs: Why we
794 should stop ignoring them. Bioessays 34: 1025-1034 doi:
795 10.1002/bies.201200073
796 Bradford MM (1976) Rapid and sensitive method for quantitation of microgram
797 quantities of protein utilizing principle of protein-dye binding. Anal Biochem 72:
798 248−254 doi: 10.1006/abio.1976.9999
799 Brown SH, Bashkirova L, Berka R, Chandler T, Doty T, McCall K, McCulloch M,
800 McFarland S, Thompson S, Yaver D, Berry A (2013) Metabolic engineering of
801 *Aspergillus oryzae* NRRL 3488 for increased production of L-malic acid. Appl
802 Microbiol Biotechnol 97: 8903−8912 doi: 10.1007/s00253-013-5132-2
803 Cheng CK, Au CH, Wilke SK, Stajich JE, Zolan ME, Pukkila PJ, Kwan HS (2013) 5′-
804 Serial Analysis of Gene Expression studies reveal a transcriptomic switch during
805 fruiting body development in *Coprinopsis cinerea*. BMC Genomics 14: 17 doi:
806 10.1186/1471-2164-14-195
807 Davuluri RV, Suzuki Y, Sugano S, Plass C, Huang THM (2008) The functional
808 consequences of alternative promoter use in mammalian genomes. Trends Genet
809 24: 167−177 doi: 10.1016/j.tig.2008.01.008
810 Forrest ARR, Kawaji H, Rehli M, Baillie JK, de Hoon MJL, Haberle V, Lassmann T,

Kulakovskiy IV, Lizio M, Itoh M, et al. (2014) A promoter-level mammalian expression atlas. Nature 507: 462-+ doi: 10.1038/nature13182

Fujii R, Ugai T, Ichinose H, Hatakeyama M, Kosaki T, Gomi K, Fujii I, Minami A, Oikawa H (2016) Reconstitution of biosynthetic machinery of fungal polyketides: Unexpected oxidations of biosynthetic intermediates by expression host. Biosci Biotechnol Biochem 80:426−431 doi: 10.1080/09168451.2015.1104234

Fujioka T, Mizutani O, Furukawa K, Sato N, Yoshimi A, Yamagata Y, Nakajima T, Abe K (2007) MpkA-dependent and -independent cell wall integrity signaling in *Aspergillus nidulans*. Eukaryot Cell 6: 1497−1510 doi: 10.1128/ec.00281-06

Gibbons JG, Salichos L, Slot JC, Rinker DC, McGary KL, King JG, Klich MA, Tabb DL, McDonald WH, Rokas A (2012) The evolutionary imprint of domestication on genome variation and function of the filamentous fungus *Aspergillus oryzae*. Curr Biol 22:1403–1409 doi: 10.1016/j.cub.2012.05.033.

Gibbons JG, Rokas A (2013) The function and evolution of the *Aspergillus* genome. Trends Microbiol 21: 14–22. doi:10.1016/j.tim.2012.09.005.

Goebels C, Thonn A, Gonzalez-Hilarion S, Rolland O, Moyrand F, Beilharz TH, Janbon G (2013) Introns regulate gene expression in Cryptococcus neoformans in a Pab2p dependent pathway. PLoS Genet 9: 15 doi: 10.1371/journal.pgen.1003686

Gomi K (2019) Regulatory mechanisms for amylolytic gene expression in the koji mold *Aspergillus oryzae*. Biosci Biotechnol Biochem 83: 1385−1401. doi:10.1080/09168451.2019.1625265

Gomi K, Iimura Y, Hara S (1987) Integrative transformation of *Aspergillus oryzae* with a plasmid containing the *Aspergillus nidulans argB* gene. Agric Biol Chem 51: 2549−2555 doi: 10.1271/bbb1961.51.2549.

Guo N, Qian Y, Zhang QQ, Chen XX, Zeng GH, Zhang X, Mi WB, Xu C, Leger RJS, Fang WG (2017) Alternative transcription start site selection in Mr-OPY2 controls lifestyle transitions in the fungus *Metarhizium robertsii*. Nat Commun 8: 13 doi: 10.1038/s41467-017-01756-1

Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol 166: 557−580 doi: 10.1016/s0022-2836(83)80284-8

Hashimoto S, Suzuki Y, Kasai Y, Morohoshi K, Yamada T, Sese J, Morishita S, Sugano S, Matsushima K (2004) 5′-end SAGE for the analysis of transcriptional start sites. Nat Biotechnol 22: 1146−1149 doi: 10.1038/nbt998

Holland MJ, Holland JP (1978) Isolation and identification of yeast messenger ribonucleic acids coding for enolase, glyceraldehyde-3-phosphate dehydrogenase, and phosphoglycerate kinase. Biochemistry 17: 4900−4907 doi:

10.1021/bi00616a007
Hunter AJ, Jin B, Kelly JM (2011) Independent duplications of α-amylase in different strains of *Aspergillus oryzae*. Fungal Genet Biol 48:438–444 doi: 10.1016/j.fgb.2011.01.006.
Hynes MJ, Szewczyk E, Murray SL, Suzuki Y, Davis MA, Lewis HMS (2007) Transcriptional control of gluconeogenesis in *Aspergillus nidulans*. Genetics 176: 139−150 doi: 10.1534/genetics.107.070904
Inoue H, Nojima H, Okayama H (1990) High-efficiency transformation of *Escherichia coli* with plasmids. Gene 96: 23−28 doi: 10.1016/0378-1119(90)90336-p
Jefferson RA (1989) The GUS reporter gene system. Nature 342: 837−838 doi: 10.1038/342837a0
Kaur JN, Panepinto JC (2016) Morphotype-specific effector functions of *Cryptococcus neoformans* PUM1. Sci Rep 6: 9 doi: 10.1038/srep23638
Li H, Hou JY, Bai L, Hu CS, Tong P, Kang YN, Zhao XD, Shao ZF (2015) Genome-wide analysis of core promoter structures in *Schizosaccharomyces pombe* with DeepCAGE. RNA Biol 12: 525−537 doi: 10.1080/15476286.2015.1022704
Liu C, Tagami K, Minami A, Matsumoto T, Frisvad JC, Suzuki H, Ishikawa J, Gomi K, Oikawa H (2015) Reconstitution of biosynthetic machinery for the synthesis of the highly elaborated indole diterpene penitrem. Angew Chem Int Ed 54:5748−5752 doi: 10.1002/anie.201501072
Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. Methods 25: 402−408 doi: 10.1006/meth.2001.1262
Machida M, Asai K, Sano M, Tanaka T, Kumagai T, Terai G, Kusumoto KI, Arima T, Akita O, Kashiwagi Y, et al. (2005) Genome sequencing and analysis of *Aspergillus oryzae*. Nature 438: 1157−1161 doi: 10.1038/nature04300
Machida M, Chang YC, Manabe M, Yasukawa M, Kunihiro S, Jigami Y (1996) Molecular cloning of a cDNA encoding enolase from the filamentous fungus, *Aspergillus oryzae*. Curr Genet 30: 423−431 doi: 10.1007/s002940050152
Machida M, Yamada O, Gomi K (2008) Genomics of *Aspergillus oryzae*: Learning from the history of koji mold and exploration of its future. DNA Res 15: 173−183 doi: 10.1093/dnares/dsn020
Maeda H, Sano M, Maruyama Y, Tanno T, Akao T, Totsuka Y, Endo M, Sakurada R, Yamagata Y, Machida M, et al. (2004) Transcriptional analysis of genes for energy catabolism and hydrolytic enzymes in the filamentous fungus *Aspergillus oryzae* using cDNA microarrays and expressed sequence tags. Appl Microbiol

883 Biotechnol 65: 74−83 doi: 10.1007/s00253-004-1608-4
884 Minetoki T, Nunokawa Y, Gomi K, Kitamoto K, Kumagai C, Tamura G (1996) Deletion
885 analysis of promoter elements of the *Aspergillus oryzae agdA* encoding α-
886 glucosidase. Curr Genet 30: 432-438 doi: 10.1007/s002940050153
887 Miura F, Kawaguchi N, Sese J, Toyoda A, Hattori M, Morishita S, Ito T (2006) A large-
888 scale full-length cDNA analysis to explore the budding yeast transcriptome. Proc
889 Natl Acad Sci U S A 103: 17846−17851 doi: 10.1073/pnas.0605645103
890 Mizutani O, Kudo Y, Saito A, Matsuura T, Inoue H, Abe K, Gomi K (2008) A defect of
891 *ligD* (human lig4 homolog) for nonhomologous end joining significantly
892 improves efficiency of gene-targeting in *Aspergillus oryzae*. Fungal Genet Biol
893 45: 878−889 doi: 10.1016/j.fgb.2007.12.010
894 Morris DR, Geballe AP (2000) Upstream open reading frames as regulators of mRNA
895 translation. Mol Cell Biol 20: 8635−8642 doi: 10.1128/mcb.20.23.8635-
896 8642.2000
897 Nakajima K, Kunihiro S, Sano M, Zhang Y, Eto S, Chang YC, Suzuki T, Jigami Y,
898 Machida M (2000) Comprehensive cloning and expression analysis of glycolytic
899 genes from the filamentous fungus, *Aspergillus oryzae*. Curr Genet 37: 322−327
900 doi: 10.1007/s002940050534
901 Oda K, Kakizono D, Yamada O, Iefuji H, Akita O, Iwashita K (2006) Proteomic analysis
902 of extracellular proteins from *Aspergillus oryzae* grown under submerged and
903 solid-state culture conditions. Appl Environ Microbiol 72: 3448−3457 doi:
904 10.1128/aem.72.5.3448-3457.2006
905 Payne GA, Nierman WC, Wortman JR, Pritchard BL, Brown D, Dean RA, Bhatnagar D,
906 Cleveland TE, Machida M, Yu J (2006) Whole genome comparison of *Aspergillus*
907 *flavus* and *A. oryzae*. Med Mycol 44:S9–S11 doi: 10.1080/13693780600835716.
908 Prade RA, Timberlake WE (1993) The *Aspergillus nidulans brlA* regulatory locus
909 consists of overlapping transcription units that are individually required for
910 conidiophore development. EMBO J 12: 2439−2447 doi: 10.1002/j.1460-
911 2075.1993.tb05898.x
912 Rojas-Duran MF, Gilbert WV (2012) Alternative transcription start site selection leads to
913 large differences in translation activity in yeast. RNA 18: 2299−2305 doi:
914 10.1261/rna.035865.112
915 Rose AB, Emami S, Bradnam K, Korf I (2011) Evidence for a DNA-based mechanism of
916 intron-mediated enhancement. Front Plant Sci 2: 9 doi: 10.3389/fpls.2011.00098
917 Roumelioti K, Vangelatos I, Sophianopoulou V (2010) A cryptic role of a glycolytic-
918 gluconeogenic enzyme (aldolase) in amino acid transporter turnover in

Aspergillus nidulans. Fungal Genet Biol 47: 254−267 doi: 10.1016/j.fgb.2009.12.004

Sakai K, Kinoshita H, Nihira T (2012) Heterologous expression system in *Aspergillus oryzae* for fungal biosynthetic gene clusters of secondary metabolites. Appl Microbiol Biotechnol 93: 2011−2022 doi: 10.1007/s00253-011-3657-9

Sato A, Oshima K, Noguchi H, Ogawa M, Takahashi T, Oguma T, Koyama Y, Itoh T, Hattori M, Hanya Y (2011) Draft genome sequencing and comparative analysis of *Aspergillus sojae* NBRC4239. DNA Res 18:165–176 doi: 10.1093/dnares/dsr009.

Shiraki T, Kondo S, Katayama S, Waki K, Kasukawa T, Kawaji H, Kodzius R, Watahiki A, Nakamura M, Arakawa T, et al. (2003) Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. Proc Natl Acad Sci U S A 100: 15776−15781 doi: 10.1073/pnas.2136655100

Sibthorp C, Wu HH, Cowley G, Wong PWH, Palaima P, Morozov IY, Weedall GD, Caddick MX (2013) Transcriptome analysis of the filamentous fungus *Aspergillus nidulans* directed to the global identification of promoters. BMC Genomics 14: 18 doi: 10.1186/1471-2164-14-847

Suzuki Y, Murray SL, Wong KH, Davis MA, Hynes MJ (2012) Reprogramming of carbon metabolism by the transcriptional activators AcuK and AcuM in *Aspergillus nidulans*. Mol Microbiol 84: 942−964 doi: 10.1111/j.1365-2958.2012.08067.x

Tada S, Gomi K, Kitamoto K, Takahashi K, Tamura G, Hara S (1991) Construction of a fusion gene comprising the Taka-amylase A promoter and the *Escherichia coli* β-glucuronidase gene and analysis of its expression in *Aspergillus oryzae*. Mol Gen Genet 229: 301−306 doi: 10.1007/bf00272170

Tagami K, Liu C, Minami A, Noike M, Isaka T, Fueki S, Shichijo Y, Toshima H, Gomi K, Dairi T, Oikawa H (2013) Reconstitution of biosynthetic machinery for indole-diterpene paxilline in *Aspergillus oryzae*. J Am Chem Soc 135:1260−1263 doi: 10.1021/ja3116636.

Taggart J, MacDiarmid CW, Haws S, Eide DJ (2017) Zap1-dependent transcription from an alternative upstream promoter controls translation of RTC4 mRNA in zinc-deficient *Saccharomyces cerevisiae*. Mol Microbiol 106: 678−689 doi: 10.1111/mmi.13851

Tanaka M, Gomi K (2013) Strategies for increasing the production level of heterologous proteins in *Aspergillus oryzae*. In Anazawa H, Shimizu S (ed) Microbial production: From genome design to cell engineering, Springer Japan, Tokyo,

pp.149−164 doi: 10.1007/978-4-431-54607-8_14

Tanaka M, Tokuoka M, Shintani T, Gomi K (2012) Transcripts of a heterologous gene encoding mite allergen Der f 7 are stabilized by codon optimization in *Aspergillus oryzae*. Appl Microbiol Biotechnol 96: 1275−1282 doi: 10.1007/s00253-012-4169-y

Toda T, Sano M, Honda M, Rimoldi O, Yang Y, Yamamoto M, Takase K, Hirozumi K, Kitamoto K, Minetoki T, et al. (2001) Deletion analysis of the enolase gene (*enoA*) promoter from the filamentous fungus *Aspergillus oryzae*. Curr Genet 40: 260−267 doi: 10.1007/s00294-001-0258-7

Tresenrider A, Ünal E (2018) One-two punch mechanism of gene repression: a fresh perspective on gene regulation. Curr Genet 64: 581−588 doi: 10.1007/s00294-017-0793-5.

Tsuboi H, Koda A, Toda T, Minetoki T, Hirotsune M, Machida M (2005) Improvement of the *Aspergillus oryzae* enolase promoter (P-*enoA*) by the introduction of *cis*-element repeats. Biosci Biotechnol Biochem 69: 206−208 doi: 10.1271/bbb.69.206

Wakai S, Yoshie T, Asai-Nakashima N, Yamada R, Ogino C, Tsutsumi H, Hata Y, Kondo A (2014) L-lactic acid production from starch by simultaneous saccharification and fermentation in a genetically engineered *Aspergillus oryzae* pure culture. Bioresour Technol 173: 376–383. doi: 10.1016/j.biortech.2014.09.094

Villar D, Berthelot C, Aldridge S, Rayner TF, Lukk M, Pignatelli M, Park TJ, Deaville R, Erichsen JT, Jasinska AJ, et al. (2015) Enhancer evolution across 20 mammalian species. Cell 160: 554−566 doi: 10.1016/j.cell.2015.01.006

Yamada O, Lee BR, Gomi K (1997) Transformation system for *Aspergillus oryzae* with double auxotrophic mutations, *niaD* and *sC*. Biosci Biotechnol Biochem 61: 1367−1369 doi: 10.1271/bbb.61.1367

Yang L, Lubeck M, Lubeck PS (2017) *Aspergillus* as a versatile cell factory for organic acid production. Fungal Biol Rev 31: 33−49 doi: 10.1016/j.fbr.2016.11.001

Yoshimi A, Yamaguchi S, Fujioka T, Kawai K, Gomi K, Machida M, Abe K (2018) Heterologous production of a novel cyclic peptide compound, KK-1, in *Aspergillus oryzae*. Front Microbiol 9: 12 doi: 10.3389/fmicb.2018.00690

Figure legends:

Fig. 1 *enoA* alternative TSS use depends on carbon source.

a *enoA* 5′ end clones obtained by 5′ SAGE. Total RNA samples were prepared from mycelia grown in YPD liquid medium for 24 h or wheat bran solid medium for 30 h. The two RNA samples were mixed and used for 5′ SAGE. Numbers represent the 5′ end peak location (arrows).

b Schematic representation of two TSSs in *enoA* and the primer binding sites used for qRT-PCR analysis. Numbers represent the positions of each TSS relative to the adenine nucleotide of the start codon (+1). The two primer sets shown in red discriminate between the two *enoA* transcript variants, because each Fw primer anneals to a distinct TSS-derived 5′ UTR. The primers shown in black were used to obtain the control signal for normalization. qRT-PCR using these primer sets could estimate the TSS-derived transcript levels relative to the total *enoA* transcript.

c uTSS or dTSS-derived *enoA* transcript levels depending on the carbon source. Total RNA samples were prepared from mycelia grown in MM containing 2% glucose for 36 h and then transferred to fresh MM without a carbon source or containing 2% glucose, fructose, glycerol, sodium acetate, or ethanol for 4 h. Values represent the means of three independent experiments. Error bars represent the standard errors

d *enoA* RT-PCR analysis. The upstream Fw primer and Rv primer of the primer sets for normalization shown in Fig. 1B were used.

e Top panel: Northern blot analysis on *A. oryzae enoA*. Total RNA samples were prepared from mycelia grown in MM containing 2% glucose for 36 h and then transferred to fresh MM without a carbon source or containing 2% glucose, fructose, glycerol, sodium acetate,

or ethanol for 4 h. Bottom panel: *enoA* transcript level quantification. Transcript signal intensity was normalized using the 18S rRNA signal. The amount of *enoA* transcript in the presence of glucose was set to 1.0. Values represent the means of three independent experiments. Error bars represent the standard errors.

Fig. 2 Highly conserved sequences in enolase-encoding gene promoters among *Aspergilli*

a Two highly conserved sequences within enolase-encoding gene promoters in *Aspergilli*. Motif identification was performed using MEME software on four 5′ sequences of 1,000 bp from the start codon in four *Aspergillus* species: *A. niger*, *A. nidulans*, *A. fumigatus*, and *A. oryzae*.

b Schematic representation of the CE_1 and CE_2 positions in the *A. oryzae enoA* promoter. Numbers represent the CE_1 and CE_2 positions relative to the adenine nucleotide of the start codon (+1).

c The CE_1 sequence in the *A. oryzae enoA* promoter and mutations used for the GUS reporter assay. The AcuK/AcuM binding motif in the CE_1 sequence is bold and underlined. The base substitutions are shown in red.

d *enoA* transcription levels in wild-type, Δ*acuK*, and Δ*acuM* strains in the presence of acetate. The primer set used for amplifying the uTSS-derived *enoA* transcript is shown in Fig. 1B. The amount of the *enoA* transcript was normalized to *histone H4*. Total RNA samples were prepared from mycelia grown in MM containing 2% glucose for 36 h and then transferred to fresh MM with 2% sodium acetate for 4 h. Values represent the means of three independent experiments. Error bars represent the standard errors. *P*-values were calculated by Student's *t*-test. *: $P < 0.05$ versus WT.

e and **f** GUS activity of the transformants harboring GUS gene (*uidA*) expression constructs in glucose (**e**) or acetate (**f**) culture conditions. *uidA* was expressed by the *enoA* promoter with or without mutations in the AcuK/AcuM binding motif. Mycelia were grown in MM containing 2% glucose for 36 h or 2% sodium acetate for 72 h. Values represent the means of three independent experiments. Error bars represent the standard errors. *P*-values were calculated by Student's *t*-test. *: $P < 0.05$ versus WT. ns: not significant.

g The CE_2 sequence in the *A. oryzae enoA* promoter and mutations used for GUS reporter assays. The conserved nucleotides are represented in bold and the base substitutions of the mutations are shown in bold red.

h and **i** Activity of transformants harboring GUS gene (*uidA*) expression constructs in glucose (**h**) or acetate (**i**) culture conditions. *uidA* was expressed by the *enoA* promoter with or without mutations at the CE_2 sequence shown in Fig. 2G. Mycelia samples were cultivated in MM containing 2% glucose for 36 h or 2% sodium acetate for 72 h. Values represent the means of three independent experiments. Error bars represent the standard errors. *P*-values were calculated by unpaired Student's *t*-test. *: $P < 0.05$, **: $P < 0.01$ versus WT. ns: not significant. n.d.: not detected.

Fig. 3 Functional analysis of the 5′ UTR intron in the *enoA* promoter

a Schematic representation of the 5′ UTR intron in the *A. oryzae enoA* promoter and mutations used for the GUS reporter assay. Intron sequence is shown in blue. 5′ and 3′ splice sites of the intron sequence are represented in bold blue and the base substitution of mutations is shown in bold red. Δi indicates the intron deletion mutation.

b and **c** GUS activity of the transformants harboring the GUS gene (*uidA*) expression

constructs in glucose (**b**) or acetate (**c**) culture conditions. *uidA* was expressed by the *enoA* promoter with or without mutations in the 5′ UTR intron. Mycelia were grown in MM containing 2% glucose for 48 h or 2% sodium acetate for 72 h. Values are the means of three independent experiments. Error bars represent the standard errors. *P*-values were calculated by Student's *t*-test. *: $P < 0.05$, **: $P < 0.01$ versus WT.

d Northern blot analysis of *uidA*. Total RNA samples were prepared from mycelia grown in MM containing 2% glucose for 36 h and then transferred to fresh MM without a carbon source or containing 2% glucose or sodium acetate and incubated for 4 h.

e GUS activity of the transformants harboring *uidA* expression constructs in acetate culture condition. *uidA* was expressed by the *enoA* promoter with or without mutations in the 5′ UTR intron and in the AcuK/AcuM binding motif. Mycelia were grown in MM containing 2% sodium acetate for 72 h. Values are the means of three independent experiments. Error bars represent the standard errors. *P*-values were calculated by Student's *t*-test. **: $P < 0.01$.

Fig. 4 Physiological significance of alternative TSS usage in *enoA*

a Schematic representation of the replacement of the native *enoA* promoter with the promoter with or without the mCS3 mutation shown in Fig. 2G or 5′ ssm mutation shown in Fig. 4A. Position of each mutation is shown by a red dot.

b Growth phenotypes of the strains on agar plates with glucose or acetate. Conidia (10^3) were inoculated on agar plates of MM containing 1% glucose or 1% sodium acetate for 4 days at 30□.

Fig. 5 TSS characterization in additional glycolytic and gluconeogenic genes in *A.*

oryzae

a Top panel: The number of *gpdA* 5′ end clones obtained by 5′ SAGE analysis as described in Fig. 1A. Numbers represent the distance of the 5′ end peak (arrows). Bottom panel: Schematic representation of two *gpdA* TSSs. Numbers represent the positions of each TSS and the exon/intron junctions relative to the adenine nucleotide of the start codon (+1).

b uTSS or dTSS-derived *gpdA* transcript levels depending on the carbon source species. Primer sets were designed using the same strategy as in Fig. 1B. Total RNA samples were prepared from mycelia grown in MM containing 2% glucose for 36 h or containing 2% sodium acetate for 72 h. Values represent the means of three independent experiments.

c Top panel: The number of *fbaA* 5′ ends obtained by 5′ RACE. Ten to twelve clones were obtained from each total RNA sample using RLM-RACE. All clones were sequenced. Total RNA samples were prepared from mycelia grown in MM containing 2% glucose for 36 h or containing 2% sodium acetate for 72 h. The clones of 5′ ends obtained from samples in glucose and acetate culture conditions are shown in blue and in red, respectively. White and black arrowheads represent the uTSS and dTSS, respectively. Bottom panel: Schematic representation of 5′ end transcripts obtained by 5′ RACE. Numbers represent the locations of exon/intron junctions within the 5′ UTR relative to the adenine nucleotide of the start codon (+1).

d Top panel: The number of *pycA* 5′ ends obtained by 5′ RACE as described in Fig. 3C. White and black arrowheads represent the uTSS and dTSS, respectively. Hatched arrowheads indicate an additional TSS within an intron in the 5′ UTR. Bottom panel: Schematic representation of 5′ end transcripts obtained by 5′ RACE. Numbers represent the locations of exon/intron junctions within 5′ UTR relative to the adenine nucleotide of

the start codon (+1).

e Classification of the TSS types in *A. oryzae* glycolytic/gluconeogenic genes. Manually-annotated genes involved in glycolysis and gluconeogenesis are bold italicized. Genes that possess alternative TSSs are shown in red. Genes possessing one TSS are shown in black. Genes that were not tested are shown in grey.

Fig. 6 Characterization of distinct TSS usage in response to carbon source in *A. nidulans acuN*

a Top panels: The number of *acuN* 5′ ends obtained by 5′ RACE. Eleven to thirteen 5′ end clones were obtained from each total RNA sample using RLM-RACE. All clones were sequenced. Total RNA samples were prepared from mycelia of *A. niduans* FGSC A4 strain grown in MM containing 2% glucose for 36 h and then transferred to fresh MM containing 2% glucose or 2% sodium acetate. The 5′ end clones obtained from samples in glucose and acetate culture conditions are shown in blue and in red, respectively. White and black arrowheads represent the uTSS and dTSS, respectively. Bottom panel: Schematic representation of 5′ end transcripts obtained by 5′ RACE. Numbers represent the locations of the highly conserved sequences, CE_1 and CE_2, and exon/intron junctions within 5′ UTR relative to the adenine nucleotide of the start codon (+1).

b uTSS- or dTSS-derived *acuN* transcript depending on the carbon source species. Primer sets were designed using the same strategy as in Fig. 1B. Total RNA samples were prepared from mycelia grown in MM containing 2% glucose for 36 h and then transferred to fresh MM without carbon source or containing 2% glucose, fructose, glycerol, sodium acetate, or ethanol for 4 h. Values represent the means of three independent experiments. Error bars represent the standard errors.

c Top panel: Northern blot analysis of *A. nidulans acuN*. Total RNA samples were prepared from mycelia grown in MM containing 2% glucose for 36 h and then transferred to fresh MM without a carbon source or media containing 2% glucose, fructose, glycerol, sodium acetate, or ethanol for 4 h. Bottom panel: Quantification of *acuN* transcript levels. The *acuN* transcript signal intensity was normalized to the 18S rRNA signal. The amount of *acuN* transcript in glucose conditions was set to 1.0. Values represent the means of three independent experiments. Error bars represent the standard errors.

d Total transcript levels and the usage pattern of alternative TSSs in *A. oryzae enoA* (top panel) and *A. nidulans acuN* (bottom panel) under glycolytic and gluconeogenic conditions. The ratio of total transcript level in *enoA* and *acuN* is same as Fig. 1E and Fig. 6C, respectively. The ratio of each transcript derived from each TSSs in total *enoA* and *acuN* transcripts was estimated from the qRT-PCR results in Fig. 1B and Fig. 6B, respectively.