Original Research Article

Identification of neutral genome integration sites with high expression and high integration efficiency in *Fusarium venenatum* TB01

Sheng Tong [a,b,**,1], Kexin An [a,b,1], Wuxi Chen [a,b], Mengdan Chai [a,b], Yuanxia Sun [a,b], Qinhong Wang [a,b], Demao Li [a,b,*]

[a] *Tianjin Key Laboratory for Industrial Biological Systems and Bioprocessing Engineering, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308, China*
[b] *National Innovation Centre for Synthetic Biology, Tianjin, 300308, China*

ABSTRACT

CRISPR/Cas9-mediated homology-directed recombination is an efficient method to express target genes. Based on the above method, providing ideal neutral integration sites can ensure the reliable, stable, and high expression of target genes. In this study, we obtained a fluorescent transformant with neutral integration and high expression of the *GFP* expression cassette from the constructed *GFP* expression library and named strain FS. The integration site mapped at 4886 bp upstream of the gene FVRRES_00686 was identified in strain FS based on a Y-shaped adaptor-dependent extension, and the sequence containing 600 bp upstream and downstream of this site was selected as the candidate region for designing sgRNAs (Sites) for CRISPR/Cas9-mediated homology-directed recombination. PCR analysis showed that the integration efficiency of CRISPR/Cas9-mediated integration of target genes in designed sites reached 100%. Further expression stability and applicability analysis revealed that the integration of the target gene into the above designed sites can be stably inherited and expressed and has no negative effect on the growth of *F. venenatum* TB01. These results indicate the above designed neutral sites have the potential to accelerate the development of *F. venenatum* TB01 through overexpression of target genes in metabolic engineering.

1. Introduction

With the increase in the global population and protein demand, the daily protein supply depended on traditional animal husbandry and agriculture is facing numerous challenges [[1–3]]. Moreover, considering that traditional animal husbandry relies on water and land resources and is a major source of global greenhouse gas emissions [4,5], a new pattern of protein supply for consumption is urgently needed. *Fusarium venenatum*, discovered from more than 3000 soil organism samples, has been successfully cultured as a mycoprotein source under the trade name of "Quorn" [[6–9]]. Mycoproteins with a fibrous texture

have high security [[10–12]] and offer a good nutritional balance, including zero cholesterol, high dietary fiber, and richness in essential amino acids [13]. Therefore, it has the potential to partially or entirely substitute animal- and plant-derived protein in the human diet.

To further promote the development of *F. venenatum* through genetic engineering, efficient genetic manipulation tools are necessary. Recently, a classical gene deletion method (homologous recombination) based on homology arms [14] and an efficient DNA-free CRISPR/Cas9 system based on the AMA1-based *Cas9* expression vector [15] were established by our group in *F. venenatum* TB01. Using the above system, the knockout or deletion of target genes can be successfully

accomplished. However, the means to ensure the reliable, efficient, and stable expression of target genes still need to be explored.

Currently, the expression of target genes in fungi mainly depends on episomal vectors and genome integration. The episomal vectors can replicate autonomously in the host based on a self-replicator such as AMA1 and CEN/ARS [16,17]; however, the obvious disadvantage is that they are readily lost without selective pressure, including in *F. venenatum* TB01 [[15,18–22]]. Compared to self-replicating vectors, integration of target genes into the genome for overexpression is a more stable strategy [22]. However, because of the nonhomologous end joining (NHEJ) pathway of DNA double strand break (DSB) repair, the integration site of foreign DNA fragments in the genome is random, resulting in various expression levels of target genes [23]. In addition, NHEJ-mediated random integration may disrupt important endogenous genes and cause some new traits unrelated to the target gene [22]. The site-specific integration method based on neutral genome integration sites for gene overexpression has been successfully applied in microbial genetic engineering [[24–27]]; however, to date, similar studies have not been reported in *F. venenatum* TB01. Moreover, considering that NHEJ is the dominant pathway for repairing DSBs in *F. venenatum* TB01 [14], neutral genome integration sites with high expression and high integration efficiency used for CRISPR/Cas9-mediated homology-directed recombination of target genes urgently need to be identified.

In this study, the new neutral genome integration sites that could ensure high and stable expression and high integration efficiency of target genes based on the CRISPR/Cas9 system were obtained in *F. venenatum* TB01 by constructing a random *GFP* expression library, Y-shaped adaptor-dependent extension and characterizing the potential integration site. The integration sites identified here provide an important synthetic tool for accelerating the development of *F. venenatum* TB01 through metabolic engineering in the future.

2. Materials and methods

2.1. Strain and culture conditions

F. venenatum TB01 isolated by the Tianjin Institute of Industrial Biotechnology was deposited at CGMCC under accession number CGMCC NO. 20740. It and its derived strains expressing green fluorescent protein (GFP) were conserved as a mixture of dry conidia at −80 °C, and routinely grown in/on glucose yeast extract broth/agar (GYB/GYA) [14] at 28 °C. *Escherichia coli* DH5α (Biomed, Beijing, China) was grown in/on LB broth/agar at 37 °C for routine gene cloning and vector construction.

2.2. Vectors construction

To construct the *GFP* expression library in *F. venenatum* TB01, the promoter sequence of *tef* (FVRRES_13282) and the fragment containing the ORF of *GFP* and terminator T*trpc* were amplified with the corresponding primer pairs Ptef-1/2 and GFPT-1/2, respectively. The obtained promoter P*tef* and *GFP*-T*trpc* fragments were cloned into the *Eco*RI site of the pK2-*hpt* vector by seamless cloning (Vazyme, Nanjing, China) to generate pK2-P*tef-GFP*-T*trpc-hpt*. The *GFP* expression cassette used for integration into the genome was finally amplified from this vector with the primer pair pK-1/2 located in the vector backbone.

For construction of the CRISPR/Cas9 expression vectors, an sgRNA scaffold was synthesized by Tsingke Biotechnology Co., Ltd. (Tsingke, Beijing, China), and the sgRNA target sites of the sequence containing the integration site of the *GFP* expression cassette in the genome were identified using the sgRNACas9 tool [28]. The promoter sequence of 5S rRNA (FVRRES_5S_rRNA_393) was amplified via PCR with the primer pair 5S-1/5Ssite1-2 or 5S-1/5Ssite2-2 using *F. venenatum* TB01 genomic DNA as the template. Then, the sgRNA driven by the 5S rRNA promoter was created by overlapping PCR and cloned into the *Pac*I site of the pFC332 vector, which contained *A. niger* codon-optimized *Cas9* and

AMA1 components [29].

For construction of the donor DNA with homologous arms, the cassettes P*tef-GFP*-T*trpc* and P*gpdA-GFP*-T*trpc* were amplified from the constructed vectors pK2-P*tef-GFP*-T*trpc-hpt* and pK2-P*gpdA-GFP*-T*trpc-hpt* with the corresponding primer pairs tef-1/Ttrpc-2 and gpdA-1/Ttrpc-2, respectively. Then, the homologous arms that were homogenous to the flanking region of the sgRNA target sites were amplified via PCR using *F. venenatum* TB01 genomic DNA as the template. The donor DNA with homologous arms was finally created by overlapping PCR using the corresponding amplification products. All PCR products were amplified using PrimeSTAR® Max DNA Polymerase (Takara, Kyoto, Japan). The primers used in this study are listed in Table S1, the nucleotide sequence of the *A. niger* codon-optimized *Cas9* is listed in Table S2, the nucleotide sequences of the sgRNA expression cassettes are listed in Table S3, and the nucleotide sequences of donor DNA are listed in Table S4.

2.3. Polyethylene glycol (PEG)-mediated protoplast transformation

PEG-mediated protoplast transformation of *F. venenatum* TB01 was performed as previously described [14]. Briefly, the protoplasts were prepared using 2 mg/ml driselase (Lablead, Beijing, China) and 4 mg/ml smailase (Sangon, Shanghai, China) at 32 °C and 130 rpm for 2 h. For protoplast transformation, 100 μl protoplast suspension was mixed with 25 μl SPTC and 10 μg PCR cassette or/and expression vector and incubated on ice for 40 min. Then, 1 ml of SPTC was gently added and incubated at room temperature for 30 min. After that, the mixture was combined with regeneration medium tempered to approximately 42 °C and poured into 90-mm-diameter petri dishes. After overnight incubation, the screening medium containing 15 μg/ml hygromycin was covered.

2.4. Fluorescence observation of fungal strains

Fluorescence observation of fungal strains using a hand-held fluorescent protein excitation light source (LUYOR-3415GR, Shanghai, China) was performed as previously described [14]. Briefly, the transformation plates or mycelia collected from fermentation medium were placed in a dark environment, the green fluorescent excitation light source was turned on to illuminate them, and the colony transformants or mycelia were observed or photographed through the filter. The mean fluorescence intensity of fluorescent strains (5 transformants) cultured on GYA for 2 days was quantified using ImageJ software (http://imagej.nih.gov/ij/; provided in the public domain by the National Institutes of Health, Bethesda, USA).

2.5. Identification of the integration site in the genome

The integration site of the *GFP* expression cassette in the genome was identified as previously described [30] with some modifications using a Y-shaped adaptor-dependent extension. Total genomic DNA was isolated using a plant rapid genomic DNA kit (Biomed, Beijing, China), and then the genomic DNA was enzymatically digested using the isocaudomers *Bam*HI and *Bgl*II at 37 °C overnight. The Y-shaped adaptors were attached to the two ends of the digestion product using T4 DNA ligase (Thermo, Waltham, USA) at 22 °C for 2 h. The target fragment, including the integration site of the *GFP* expression cassette in the genome, was obtained via linear amplification and exponential amplification using the above ligation mixture as a template, and Padap-1, which was located in the adaptor, and Pcassette-2/Pcassette-3, which was located in the cassette, as primer. Ultimately, the location of the integration site in the genome was determined by sequencing and aligning the target fragment, and the relevant Circos diagram was drawn using TBtools software [31]. The primers used in this study are listed in Table S1.

2.6. PCR analysis of the transformants generated by integrating the donor DNA with homologous arms via the CRISPR/Cas9 system

The simple templates used for PCR amplification were prepared as previously described [32]. Briefly, a small amount of mycelia was collected and placed into 8 μl 0.3 M NaOH. After incubating at 98 °C for 2 min, 170 μl neutralization solution (20 mM Tris-HCl (pH 8.0), 0.3 M HCl) was added, and simple templates containing fungal genomic DNA were obtained. The PCR procedure was set according to the instructions of the 2 × Rapid Taq Master Mix (Vazyme, Nanjing, China) with the primer pair S1L-1/S1R-2 or the primer pair S2L-1/S2R-2. For the cassette P*tef-GFP*-T*trpc*, the predicted size of the PCR product in the correctly inserted transformants was ~2000 bp larger than that in the wild type (control), while for the cassette P*gpdA-GFP*-T*trpc*, the predicted size of the PCR products in the correctly inserted transformants was ~3000 bp larger than that in the wild type (control). The primers used in this study are listed in Table S1.

2.7. Observation of the morphology of mycelia

Conidial suspensions (500 μL, 5×10^6 conidia/mL) derived from GYA plates were cultured in 50 mL GYB liquid medium at 28 °C and 180 rpm for 16 h. Then, the morphology of mycelia was observed by fluorescence microscopy (Leica DM5000B, Wetzlar, Germany).

2.8. Determination of dry mycelial weight

Dry mycelial weight was determined as described [15]. Briefly, conidial suspensions (300 μL, 5×10^6 conidia/mL) derived from 10 days GYA plates were cultured in 50 mL fermentation medium (glucose 60 g/L, yeast extract 0.5 g/L, $(NH_4)_2SO_4$ 6 g/L, $MgSO_4$ 1.5 g/L, KCl 0.7 g/L, Na_2SO4 0.5 g/L, KH_2PO_4 2 g/L, $CaCO_3$ 0.5 g/L) at 28 °C and 180 rpm for 4 days. Then, the mycelia were collected by filtration using a SHZ-D(III) vacuum filtration apparatus (Baoling, Shanghai, China) and washed twice with distilled water. The washed mycelia were dried in an electric thermostatic drying oven 101-00B (Lichen, Zhejiang, China) until a constant weight was reached. The experiment was performed in triplicate, and error bars represent the standard deviation of three replicates.

2.9. Determination of glucose concentration

The *F. venenatum* TB01 cultured at 28 °C in fermentation medium was centrifuged to remove the mycelia and then filtered through a 0.22 μm filter (Jinlong, Tianjin, China). This cell-free supernatant was diluted with distilled water to a suitable concentration, and the glucose concentration was measured using an M-online biochemical analyzer (Siemans, Shenzhen, China). The carbon conversion rate (g/g) was calculated according to the following formula: dry mycelial weight/ (initial glucose concentration - final glucose concentration). The experiment was performed in triplicate, and error bars represent the standard deviation of three replicates.

3. Results

3.1. Construction of a random GFP expression library in F. venenatum TB01

Some studies have reported that NHEJ-mediated random genome integration of the *GFP* expression cassette resulted in a significant difference in fluorescence intensity among the transformants with different locations of the inserted fragment [22]. Therefore, it provides an effective strategy for constructing a genome-scale trackable insertional mutagenesis library. To identify a potential neutral integration site with high expression, a *GFP* expression cassette in which the *GFP* reporter gene was driven by the endogenous moderate promoter P*tef* [14] was transformed into *F. venenatum* TB01 by PEG-mediated protoplast transformation and used to construct a *GFP* random expression library. Subsequently, four high fluorescence expression transformants were screened from primary transformation plates using a hand-held fluorescent protein excitation light source (Fig. S1), and further evaluations of morphology, biomass and carbon conversion rate were performed. The entire process from library construction to transformants validation is shown in Fig. 1.

3.2. Identification of the neutral integration site with high expression in the genome

Based on the above system, a transformant with high *GFP* expression and neutral integration was obtained and named strain FS (Fig. 2A). The result of mycelial morphology observation showed that there was no significant difference between the transformant FS and wild type
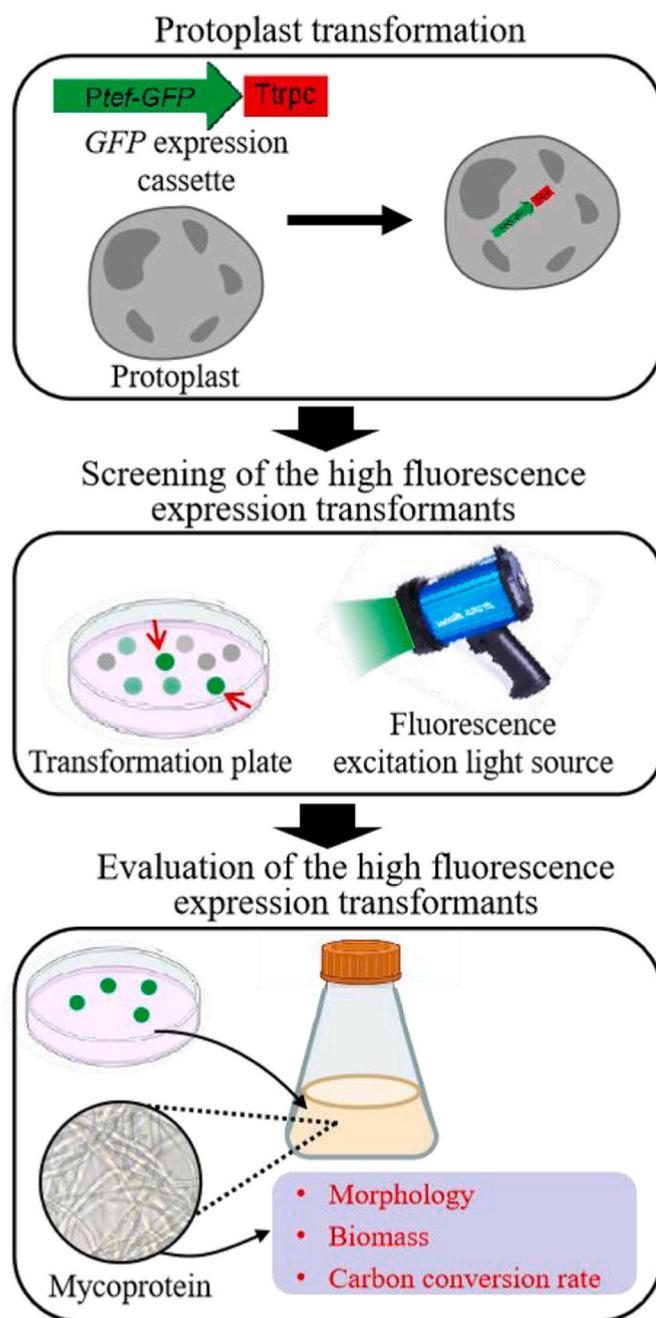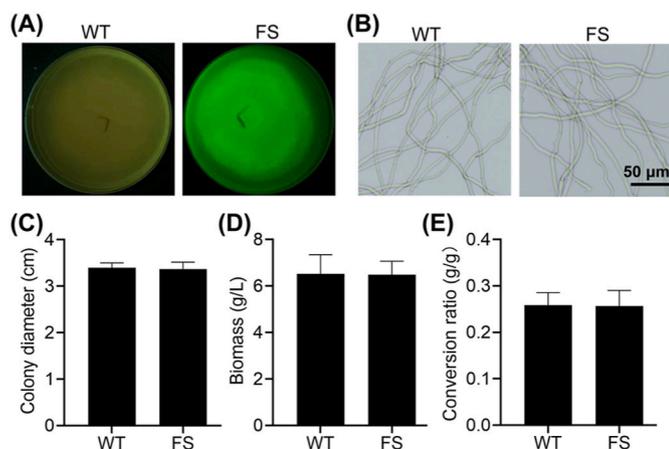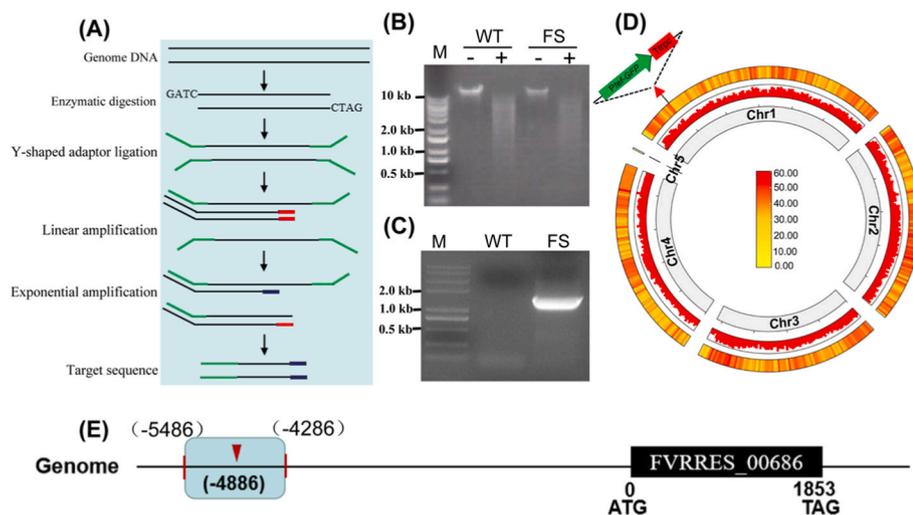

Fig. 1. Flowchart of construction of the *GFP* expression library, and screening and evaluation of the high fluorescence expression transformants.

Fig. 2. Evaluation of the high fluorescence expression transformant. (A) Representative images of fluorescent transformant FS and wild type. The colonies were observed after culturing on GYA plates at 28 °C for 5 days. (B) Hyphal morphology was observed after culturing in GYB liquid medium at 28 °C for 16 h. (C) The colony diameters of *F. venenatum* TB01 after 3 days of incubation. (D) The biomass of *F. venenatum* TB01 after 4 days of fermentation. (E) The carbon conversion ratio in *F. venenatum* TB01 after 4 days of fermentation. WT: wild type; FS: high fluorescence expression transformant. Error bars represent the standard deviation of three replicates.

(Fig. 2B). In addition, the colony diameter (3.41 vs. 3.37 cm), biomass (6.482 vs. 6.489 g/L) and carbon conversion rate (0.264 vs. 0.258 g/g) in the transformant FS were also similar to those in the wild type (Fig. 2C–E).

To identify the neutral integration site with high expression in the genome of *F. venenatum* TB01, a Y-shaped adaptor-dependent extension based on enzymatic digestion of genomic DNA was employed (Fig. 3A). As shown in Fig. 3B, after an overnight incubation at 37 °C using *Bam*HI and *Bgl*II, which produce the same cohesive end (-GATC), the genomic DNA was completely digested. Then, a target fragment including the integration site was obtained using the Y-shaped adaptor ligation mixture as the template via linear amplification and exponential amplification in order (Fig. 3C). Sequence alignment against the genome of *F. venenatum* (ASM90000737v1) showed that the integration site of the *GFP* expression cassette was located on chromosome 1 (Fig. 3D and
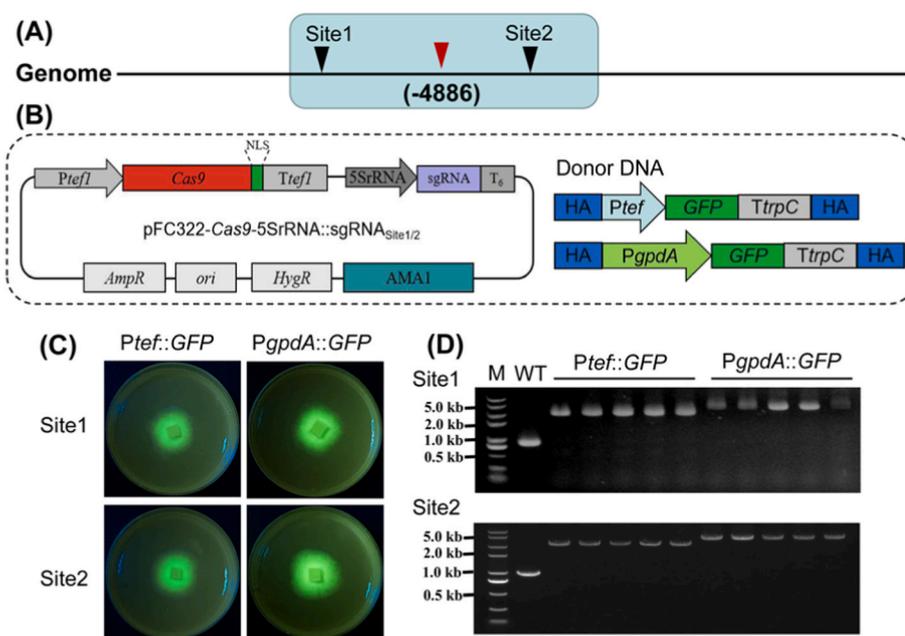
Fig. S2). Further analysis revealed that this integration site was mapped at 4886 bp upstream of gene FVRRES_00686, and the sequence containing 600 bp upstream and downstream of this site was selected as candidate region for the subsequent experiments (Fig. 3E).

3.3. Determining the integration efficiency of CRISPR/Cas9-mediated integration

To determine the integration efficiency of CRISPR/Cas9-mediated directional integration, two suitable sites from the above candidate region were identified using the sgRNACas9 tool (Fig. 4A). The availability of the corresponding sgRNAs was further confirmed by the CRISPR/Cas9 system (Fig. S3). Then, the *GFP* expression cassette driven by the endogenous promoter P*tef* or P*gpdA* was inserted into the identified sites through CRISPR/Cas9-mediated homology-directed recombination (Fig. 4B), and the corresponding fluorescent strains were verified using the hand-held fluorescent protein excitation light (Fig. 4C). To distinguish ectopic integration (band size: ~1 kb and ~3 or 4 kb) and site-specific integration (band size: only ~3 or 4 kb) in fluorescent transformants, PCR primer pairs located in the homologous arms of donor DNA were designed. PCR analysis showed that the site-specific integration efficiencies of the *GFP* expression cassettes in both identified sites reached 100% (Fig. 4D).

3.4. Evaluation of the expression stability and applicability of the two identified integration sites

To test the expression stability of the cassettes in both identified sites, the corresponding fluorescent transformants were transferred five times successively on the GYA plate. The result of fluorescence detection showed that there was no significant difference in mean fluorescence density between the first and fifth generation of fluorescent transformants regardless of being integrated with cassette P*tef*-GFP-T*trpc* or P*gpdA*-GFP-T*trpc*, and also no significant difference between the fluorescent transformants integrated with cassette P*tef*-GFP-T*trpc* and original integration transformant FS (Fig. 5A). Besides, the fluorescence signal of P*gpdA*-GFP expressing strain in fermentation medium over the time-course was consistent (Fig. S4). These above results indicated that the cassettes inserting into the identified sites can stably be inherited and high expressed. Certainly, as expected, the *GFP* expression cassette driven by promoter P*gpdA* conferred strains stronger fluorescence signal

Fig. 3. Identification of the neutral integration site in the genome of *F. venenatum* TB01. (A) Flowchart of target fragment amplification based on a Y-shaped adaptor-dependent extension. (B) Enzymatic digestion of genomic DNA using the isocaudomers *Bam*HI and *Bgl*II. WT: wild type; FS: high fluorescence expression transformant. -: untreated with restriction enzyme; +: treated with restriction enzyme. (C) PCR amplification of the target fragment including the integration site of the *GFP* expression cassette. PCR amplification of the target fragment was performed via linear amplification and exponential amplification in that order using the Y-shaped adaptor ligation mixture as a template. (D) The location of the *GFP* expression cassette in the genome. The Circos diagram was drawn by TBtools software using the genome of *F. venenatum* (ASM90000737v1). The innermost layer is the length of each chromosome, and the other two layers are the display form of gene density via bar and heatmap, respectively. (E) The integration site of the *GFP* expression cassette was mapped 4886 bp upstream of the FVRRES_00686 gene. The red triangle indicates the integration site of the *GFP* expression cassette, and the rounded rect-

angle indicates the candidate region containing 600 bp upstream and downstream of this integration site.

Fig. 4. Evaluation of the integration efficiency of CRISPR/Cas9-mediated integration. (A) Schematic representation of the distribution of sites in the genome. The red triangle indicates the integration site of the *GFP* expression cassette; the rounded rectangle indicates the candidate region containing 600 bp upstream and downstream of this integration site; the black triangle indicates the sites used for CRISPR/Cas9-mediated homology-directed recombination and designed by the sgRNACas9 tool. (B) Schematic representation of AMA1-based Cas9 and sgRNA expression vectors and corresponding donor DNA. (C) Representative images of fluorescent transformants after culturing on GYA plates at 28 °C for 2 days. (D) PCR analysis of the CRISPR/Cas9-mediated homology-directed recombination transformants. The predicted size of the PCR product in the wild type is ~1 kb and in the site-specific integration transformants is ~3 kb or ~4 kb.
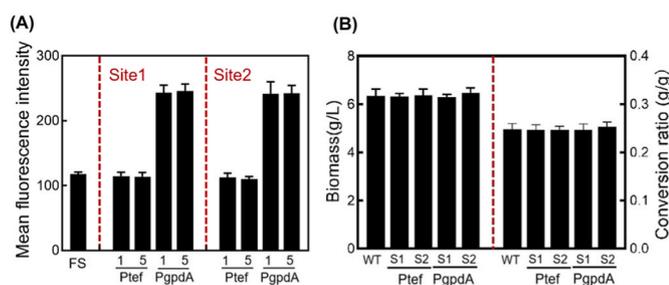
Fig. 5. Evaluation of the expression stability and applicability of the two identified integration sites. (A) Quantification of the mean fluorescence intensity of fluorescent strains by ImageJ software. (B) The biomass and carbon conversion ratio of *F. venenatum* TB01 after 4 days of fermentation. 1: first generation of fluorescent transformants; 5: fifth generation of fluorescent transformants. S1: site 1; S2: site 2. WT: wild type; P*tef*: fluorescent transformants expressing P*tef-GFP*; P*gpdA*: fluorescent transformants expressing P*gpdA-GFP*. FS: the original integration transformant. Error bars represent the standard deviation of three replicates.

than promoter P*tef* (Fig. 5A).

Given that the identified sites will be used for genetic engineering for the expression of target genes, we wanted to determine that they had no negative effect on the growth of *F. venenatum* TB01. Therefore, the biomass and carbon conversion rate between the fifth generation of fluorescent transformants and wild type were analyzed, and no significant differences were found between them (Fig. 5B). These results indicated that the identified integration sites were neutral integration sites and could be used for future genome engineering in *F. venenatum* TB01.

4. Discussion

Gene overexpression is an important part of rewriting the metabolic pathway in filamentous fungi [[33–35]]. To ensure the reliable, efficient, and stable overexpression of target genes, the ideal integration sites in the genome should be identified. Some studies have reported that NHEJ-mediated randomly inserted mutant libraries are often successfully used for the analysis of gene function and the identification of integration sites [[22,36,37]]. In this study, to facilitate the screening of

transformants with high expression of target genes for identification of integration sites, the *GFP* gene driven by the endogenous moderate promoter P*tef* [14] was used to construct a *GFP* expression library, and the transformants with high *GFP* expression were screened among colonies using the hand-held fluorescent protein excitation light (Figs. 1 and 2). Subsequently, the neutral genome integration site with high expression was successfully identified based on a Y-shaped adaptor-dependent extension and sequencing (Fig. 3).

The Y-shaped adaptor-dependent extension for identification of the integration site of the target gene has been successfully applied to plants and filamentous fungi [30,38]. Compared to other methods for identifying the integration site, such as inverse PCR and tail-PCR, Y-shaped adaptor-dependent extension has the advantages of low false-positives and high efficiency [30,38]. However, the excessively long target fragment containing the known sequence after genomic DNA digestion might lead to amplification failure when using this method. In this study, we digested the genome of *F. venenatum* TB01 with the iso-caudomers *Bam*HI and *Bgl*II, which produce the same cohesive end (-GATC), to increase the probability of obtaining the target sequence with the appropriate length and successful amplification (Fig. 3). The target sequence with the appropriate length used for attachment of the Y-shaped adaptor and amplification may be more accessible after genomic DNA treatment with a variety of flat-end enzyme mixtures. Moreover, with the development of genome sequencing technology, whole-genome resequencing is becoming an accurate and efficient way to identify the integration site of target genes in the genome.

Considering the advantages of site-specific integration method, the expression of target genes depended on the ideal neutral genome integration sites has been widely applied in animal, plant, bacteria and yeast [[24–27,39–43]]. However, far less relevant research has been conducted in filamentous fungi [[44,45]]. In this study, we successfully identified the neutral integration site of *GFP* expression cassette in the genome of *F. venenatum* TB01. From an overall perspective, the integration site was located in the region with high gene density on chromosome 1 (Fig. 3D); while from a particular perspective, the integration site was mapped at 4886 bp upstream of gene FVRRES_00686 and this 4886 bp region did not contain other genes (Fig. 3E). The special location of integration site in the genome might explain the reason that it could ensure the efficient and stable expression of *GFP* genes and had no negative effect on the growth of strain. Compared to other fungi, such as

Yarrowia lipolytica [22], the number of neutral integration sites identified in *F. venenatum* TB01 was too small. However, based on these experiences, we believe more neutral integration sites for different applications will be identified in *F. venenatum* TB01 in the future.

For homology-directed recombination of target genes at specific sites in the genome, the homologous recombination rate based on the sole homologous arm varies greatly from species to species. Even in some fungi, the recombination rate was low, which has hindered the development of strain engineering [46,47]. To improve the recombination rate of the target gene, the CRISPR/Cas9 tool and donor DNA with a homologous arm could be combined [48,49]. When using the strategy outlined above, the directional integration efficiency of the *GFP* expression cassettes in identified sites reached 100% in *F. venenatum* TB01 (Fig. 4D), which was similar to what has been reported in *Aspergillus niger* [50].

5. Conclusion

In this study, we successfully identified and verified new neutral genome integration sites with high expression used for CRISPR/Cas9-mediated homology-directed recombination of target genes in *F. venenatum* TB01. The 100% integration efficiency and high expression stability of *GFP* expression cassettes integrated into these sites indicate that they are likely to be important synthetic biological tool for accelerating the development of *F. venenatum* TB01 through metabolic engineering in the future.

Funding information

Ethical approval

This article does not concern any studies performed by any of the authors with human participants.

Data availability

The data presented in this study are available in this manuscript and the supplementary material.

CRediT authorship contribution statement

ST and DL designed the research and provide financial support. ST, KA, WC, and MC conducted most of the experiments. ST wrote the paper. ST analyzed and interpreted the data. DL, YS and QW revised the manuscript. All authors reviewed and approved the manuscript.

Declaration of competing interest

The authors declare that they have no conflicts of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.synbio.2022.12.006.

References

[1] Ritala A, Hakkinen ST, Toivari M, Wiebe MG. Single cell protein-state-of-the-art, industrial landscape and patents 2001-2016. Front Microbiol 2017;8:2009. https://doi.org/10.3389/fmicb.2017.02009.

[2] Myers SS, Smith MR, Guth S, Golden D, Vaitla B, Mueller ND, Dangour AD, Huybers P. Climate change and global food systems: potential impacts on food security and undernutrition. Annu Rev Publ Health 2017;38(1):259–77. https://doi.org/10.1146/annurev-publhealth-031816-044356.

[3] Lv X, Wu Y, Gong M, Deng J, Gu Y, Liu Y, Li J, Du G, Ledesma-Amaro R, Liu L, Chen J. Synthetic biology for future food: research progress and future directions. Future Foods 2021;3:100025. https://doi.org/10.1016/j.fufo.2021.100025.

[4] Xue Y, Luan W, Wang H, Yang Y. Environmental and economic benefits of carbon emission reduction in animal husbandry via the circular economy: case study of pig farming in Liaoning, China. J Clean Prod 2019;238(20):117968. https://doi.org/10.1016/j.jclepro.2019.117968.

[5] Lusk JL, Dan BR, Shah S, Tonsor GT. Impact of plant-based meat alternatives on cattle inventories and greenhouse gas emissions. Environ Res Lett 2022;17:024035. https://doi.org/10.1088/1748-9326/ac4fda.

[6] Wiebe MG. Myco-protein from *Fusarium venenatum*: a well-established product for human consumption. Appl Microbiol Biotechnol 2002;58:421–7. https://doi.org/10.1007/s00253-002-0931-x.

[7] Wiebe MG. Quorn™ mycoprotein-overview of a successful fungal product. Mycologist 2004;18(1):17–20. https://doi.org/10.1017/S0269-915X(04)00108-9.

[8] Finnigan TJA, Wall BT, Wilde PJ, Stephens FB, Taylor SL, Freedman MR. Mycoprotein: the future of nutritious nonmeat protein, a symposium review. Current Developments in Nutrition 2019;3(6):nzz021. https://doi.org/10.1093/cdn/nzz021.

[9] Colosimo R, Mulet-Cabero AI, Warren FJ, Edwards CH, Finnigan TJA, Wilde PJ. Mycoprotein ingredient structure reduces lipolysis and binds bile salts during simulated gastrointestinal digestion. Food Funct 2020;11:10896–906. https://doi.org/10.1039/D0FO02002H.

[10] Finnigan T, Needham L, Abbott C. Mycoprotein: a healthy new protein with a low environmental impact. In: Nadathur SR, editor. Sustainable protein sources. New York: Academic press; 2017. p. 305–25. https://doi.org/10.1016/B978-0-12-802778-3.00019-6.

[11] Jacobson MF, Deporter J. Self-reported adverse reactions associated with mycoprotein (Quorn-brand) containing foods. Ann Allergy Asthma Immunol 2018;120(6):626–30. https://doi.org/10.1016/j.anai.2018.03.020.

[12] Hashempour-Baltork F, Khosravi-Darani K, Hosseini H, Farshi P, Reihani SF. Mycoproteins as safe meat substitutes. J Clean Prod 2020;253:119958. https://doi.org/10.1016/j.jclepro.2020.119958.

[13] Thomas AB, Shetane TD, Singha RG, Nanda RK, Poddar SS, Shirsat A. Employing central composite design for evaluation of biomass production by *Fusarium venenatum*: in vivo antioxidant and antihyperlipidemic properties. Appl Biochem Biotechnol 2017;183:91–109. https://doi.org/10.1007/s12010-017-2432-5.

[14] Tong S, An K, Zhou W, Chen W, Sun Y, Wang Q, Li D. Establishment of high-efficiency screening system for gene deletion in *Fusarium venenatum* TB01. Journal of Fungi 2022;8(2):169. https://doi.org/10.3390/jof8020169.

[15] Tong S, An K, Chen W, Zhou W, Sun Y, Wang Q, Li D. Evasion of Cas9 toxicity to develop an efficient genome editing system and its application to increase ethanol yield in *Fusarium venenatum* TB01. Appl Microbiol Biotechnol 2022;106:6583–93. https://doi.org/10.1007/s00253-022-12178-5.

[16] Aleksenko A, Clutterbuck AJ. Autonomous plasmid replication in *Aspergillus nidulans*: AMA1 and MATE elements. Fungal Genet Biol 1997;21(3):373–87. https://doi.org/10.1006/fgbi.1997.0980.

[17] Vernis L, Abbas A, Chasles M, Gaillardin CM, Brun C, Huberman JA, Fournier P. An origin of replication and a centromere are both needed to establish a replicative plasmid in the yeast *Yarrowia lipolytica*. Mol Cell Biol 1997;17(4):1995–2004. https://doi.org/10.1128/mcb.17.4.1995.

[18] Katayama T, Nakamura H, Zhang Y, Pascal A, Fujii W, Maruyama J. Forced recycling of an AMA1-based genome-editing plasmid allows for efficient multiple gene deletion/integration in the industrial flamentous fungus *Aspergillus oryzae*. Appl Environ Microbiol 2019;85(3):e01896–918. https://doi.org/10.1128/AEM.01896-18.

[19] Kadooka C, Yamaguchi M, Okutsu K, Yoshizaki Y, Takamine K, Katayama T, Maruyama J, Tannaki H, Futagami T. A CRISPR/Cas9-mediated gene knockout system in *Aspergillus luchuensis* mut *kawachii*. Biosc Biotech Biochem 2020;84(10):2179–83. https://doi.org/10.1080/09168451.2020.1792761.

[20] Kralova M, Bergougnoux V, Frebort I. CRISPR/Cas9 genome editing in ergot fungus *Claviceps purpurea*. J Biotechnol 2021;325(10):341–54. https://doi.org/10.1016/j.jbiotec.2020.09.028.

[21] Garrigues S, Manzanares P, Marcos J. Application of recyclable CRISPR/Cas9 tools for targeted genome editing in the postharvest pathogenic fungi *Penicillium digitatum* and *Penicillium expansum*. Curr Genet 2022;68:515–29. https://doi.org/10.1007/s00294-022-01236-0.

[22] Liu X, Cui Z, Su T, Lu X, Hou J, Qi Q. Identification of genome integration sites for developing a CRISPR-based gene expression toolkit in *Yarrowia lipolytica*. Microb Biotechnol 2022;15:2223–34. https://doi.org/10.1111/1751-7915.14060.

[23] Bai Q, Cheng S, Zhang J, Li M, Cao Y, Yuan Y. Establishment of genomic library technology mediated by non-homologous end joining mechanism in *Yarrowia lipolytica*. Sci China Life Sci 2021;64:2114–28. https://doi.org/10.1007/s11427-020-1885-x.

[24] EauClaire SF, Zhang J, Rivera CG, Huang LL. Combinatorial metabolic pathway assembly in the yeast genome with RNA-guided Cas9. J Ind Microbiol Biotechnol 2016;43(7):1001–15. https://doi.org/10.1007/s10295-016-1776-0.

[25] Kim SK, Chung D, Himmel ME, Bomble YJ, Westpheling J. Heterologous co-expression of two β-glucanases and a cellobiose phosphorylase resulted in a significant increase in the cellulolytic activity of the *Caldicellulosiruptor bescii* exoproteome. J Ind Microbiol Biotechnol 2019;46(5):687–95. https://doi.org/10.1007/s10295-019-02150-0.

[26] Liu R, Fang F, An Z, Huang R, Wang Y, Sun X, Fu S, Fu A, Deng Z, Liu T. Genomics-driven discovery of the biosynthetic gene cluster of maduramicin and its

overproduction in *Actinomadura* sp. J1-007. J Ind Microbiol Biotechnol 2020;47 (2):275–85. https://doi.org/10.1007/s10295-019-02256-5.

[27] Yu W, Gao J, Zhai X, Zhou Y. Screening neutral sites for metabolic engineering of methylotrophic yeast *Ogataea polymorpha*. Synthetic and systems Biotechnology 2021;6(2):63–8. https://doi.org/10.1016/j.synbio.2021.03.001.

[28] Xie S, Shen B, Zhang C, Huang X, Zhang Y. SgRNAcas9: a software package for designing CRISPR sgRNA and evaluating potential of-target cleavage sites. PLoS One 2014;9(6):e100448. https://doi.org/10.1371/journal.pone.0100448.

[29] Wilson FM, Harrison RJ. CRISPR/Cas9 mediated editing of the Quorn fungus *Fusarium venenatum* A3/5 by transient expression of Cas9 and sgRNAs targeting endogenous marker gene *PKS12*. Fungal Biology and Biotechnology 2021;8:15. https://doi.org/10.1186/s40694-021-00121-8.

[30] Xiao Y, Luo M, Fang W, Luo K, Hou L, Luo X, Pei Y. PCR walking in cotton genome using YADE method. Acta Genetica Sin 2002;29:62–6 [in Chinese)].

[31] Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. Tbtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant 2020;13(8):1194–202. https://doi.org/10.1016/j.molp.2020.06.009.

[32] Tong S, Li M, Keyhani NO, Liu Y, Yuan M, Lin D, Jin D, Li X, Pei Y, Fan Y. Characterization of a fungal competition factor: production of a conidial cell-wall associated antifungal peptide. PLoS Pathog 2020;16(4):e1008518. https://doi.org/10.1371/journal.ppat.1008518.

[33] Gao L, Li Z, Xia C, Qu Y, Liu M, Yang P, Yu L, Song X. Combining manipulation of transcription factors and overexpression of the target genes to enhance lignocellulolytic enzyme production in *Penicillium oxalicum*. Biotechnol Biofuels 2017;10:100. https://doi.org/10.1186/s13068-017-0783-3.

[34] Shi T, Gao J, Wang W, Wang K, Xu G, Huang H, Ji X. CRISPR/Cas9-based genome editing in the famentous fungus *Fusarium fujikuroi* and its application in strain engineering for gibberellic acid production. ACS Synth Biol 2019;8(2):445–54. https://doi.org/10.1021/acssynbio.8b00478.

[35] Zhang S, Wakai S, Sasakura N, Tsutsumi H, Hata Y, Ogino C, Kondo A. Pyruvate metabolism redirection for biological production of commodity chemicals in aerobic fungus *Aspergillus oryzae*. Metab Eng 2020;61:225–37. https://doi.org/10.1016/j.ymben.2020.06.010.

[36] Gao F, Zhou B, Li G, Jia P, Li H, Zhao Y, Zhao P, Xia G, Guo H. A glutamic acid-rich protein identified in *Verticillium dahliae* from an insertional mutagenesis affects microsclerotial formation and pathogenicity. PLoS One 2010;5(12):e15319. https://doi.org/10.1371/journal.pone.0015319.

[37] Zhang D, Wang X, Chen J, Kong Z, Gui Y, Li N, Bao Y, Dai X. Identification and characterization of a pathogenicity-related gene *VdCYP1* from *Verticillium dahliae*. Sci Rep 2016;6:27979. https://doi.org/10.1038/srep27979.

[38] Fang W, Zhang Y, Ma J, Xiao Y, Yang X, Pei Y. Cloning and analysis of the promoter of Subtilisin-like protease gene *CDEP-1* from *Beauveria bassiana* by YADE method. Mycosystema 2003;22(2):252–8 [in Chinese)].

[39] Beard C, Hochedlinger K, Plath K, Wutz A, Jaenisch R. Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. Genesis 2010;44(1):23–8. https://doi.org/10.1002/gene.20180.

[40] Albert H, Dale EC, Lee E, Ow DW. Site-specific integration of DNA into wild-type and mutant *lox* sites placed in the plant genome. Plant J 1995;7(4):649–59. https://doi.org/10.1046/j.1365-313X.1995.7040649.x.

[41] Li Z, Xing A, Moon BP, McCardell RP, Mills K, Falco SC. Site-specific integration of transgenes in soybean via recombinase-mediated DNA cassette exchange. Plant Physiol 2009;151(3):1087–95. https://doi.org/10.1104/pp.109.137612.

[42] Hoang TT, Kutchma AJ, Becher A, Schweizer HP. Integration-proficient plasmids for *Pseudomonas aeruginosa*: site-specific integration and use for engineering of reporter and expression strains. Plasmid 2000;43(1):59–72. https://doi.org/10.1006/plas.1999.1441.

[43] Hartl B, Wehrl W, Wiegert T, Homuth G, Schumann W. Development of a new integration site within the *Bacillus subtilis* chromosome and construction of compatible expression cassettes. J Bacteriol 2001;183(8):2696–9. https://doi.org/10.1128/JB.183.14.4393-4393.2001.

[44] Chen S, Song Y, Cao J, Wang G, Wei H, Xu X, Lu L. Localization and function of calmodulin in live-cells of *Aspergillus nidulans*. Fungal Genet Biol 2010;47(3):268–78. https://doi.org/10.1016/j.fgb.2009.12.008.

[45] Roux I, Chooi YH. Cre/*lox*-mediated chromosomal integration of biosynthetic gene clusters for heterologous expression in *Aspergillus nidulans*. ACS Synth Biol 2022;11 (3):1186–95. https://doi.org/10.1021/acssynbio.1c00458.

[46] Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, Weiss RL, Borkovich KA, Dunlap JC. A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors. Proc Natl Acad Sci USA 2006;103(27):10352–7. https://doi.org/10.1073/pnas.0601456103.

[47] Shafran H, Miyara I, Eshed R, Prusky D, Sherman A. Development of new tools for studying gene function in fungi based on the gateway system. Fungal Genet Biol 2008;45(8):1147–54. https://doi.org/10.1016/j.fgb.2008.04.011.

[48] Liu Q, Gao R, Li J, Lin L, Zhao J, Sun W, Tian C. Development of a genome-editing CRISPR/Cas9 system in thermophilic fungal *Myceliophthora* species and its application to hyper-cellulase production strain engineering. Biotechnol Biofuels 2017;10:1. https://doi.org/10.1186/s13068-016-0693-9.

[49] Yu L, Xiao M, Zhu Z, Wang Y, Zhou Z, Wang P, Zou G. Efficient genome editing in *Claviceps purpurea* using a CRISPR/Cas9 ribonucleoprotein method. Synthetic and systems Biotechnology 2022;7(2):664–70. https://doi.org/10.1016/j.synbio.2022.02.002.

[50] Zheng X, Zheng P, Zhang K, Cairns TC, Meyer V, Sun J, Ma Y. 5S rRNA promoter for guide RNA expression enabled highly efficient CRISPR/Cas9 genome editing in *Aspergillus niger*. ACS Synth Biol 2018;8(7):1568–74. https://doi.org/10.1021/acssynbio.7b00456.