A new Cghd I gene with negative regulation on the synthesis of chaetoglobosin A

Cheng Jiang¹,²*, Jinzhu Song², Junzheng Zhang³

¹School of Biological Engineering, Huainan Normal University, Huainan, China
² School of Life Sciences and Technology, Harbin Institute of Technology, Harbin, China
³School of Chemical Engineering and Technology, Harbin Institute of Technology, Harbin, China

*Corresponding author e-mail: jiangcheng0613@163.com

Abstract. Chaetoglobosins have wide application prospect, but the low yields make it not widely used. The aim of this study was to obtain high-yielding strains and to find the genes regulating the synthesis of chaetoglobosins. A strain with yield increased to 115% was obtained through inserting a short-DNA randomly into the genome of Chaetomium globosum by agrobacterium. At the same time, inverse PCR was used to determine the insertion position of T-DNA. A new gene Cghd I with negative regulation on the synthesis of chaetogllobosin A was found, indicating that the detectable random insertability mutation technique established in this paper can be used to obtain high yield strains and find regulatory genes.

1. Introduction

Chaetoglobosins are a group of macrocyclic polyketides, belonging to the family of cytochalasins, which has a wide range of biological activities [1]. Chaetoglobosin A was first discovered from Chaetomium globosum and reported to have the ability to inhibit Hela cells in 1973 [2]. Since then, more than 70 members in this family have been reported, many of which have the ability to inhibit microorganism and cancer cells [3].

Polyketides compound by polyketide synthase (PKS) through acetate polymalonyl pathway [4]. The research on the regulation of PKS had been a long time, but most of them were concentrated in prokaryotes, especially Streptomyces [5]. There are relatively few reports on polyketide synthase regulations in eukaryotes, especially in filamentous fungi. Kurosaki found that NADPH in carrot and benzoyl-sensitive amino acid residues located in the catalytic centre were crucial for the synthesis of cerulenin, and the accumulation of cerulenin will also cause feedback inhibition to this system [6]. Lee knocked out the histone deacetylase gene in Aspergillus fumigatus and found that a variety of secondary metabolites were increased, but since the A. fumigatus did not synthesize polyketides, it was not known whether this gene had an inhibitory effect on the synthesis of polyketides [7]. Hong found a gene NPG1 encoding phosphopantetheinyl transferases (PPTase) from a variety of black yeast (Aureobasidium pullulans var. melanogenum), which had an activation effect on pks gene. The deletion of NPG1 will lead to the disappearance of melanin, a polyketide [8]. Tao overexpressed a gene Lae A with positive global regulation function in C. globosum, and both of a high-yielding strain and a new chaetoglobosin Z were obtained [9].
However, these attempts were not enough, and many genes regulating fungal polyketide synthesis remained unknown. In this study, a detectable random insertability mutation technique was established based on the modified inverse PCR method, and a new negative regulatory gene was found while the production of chaetoglobosin A increased. Our research provided an effective method to study regulatory mechanism and improve the yields of chaetoglobosins.

2. Material and method

2.1. Fungal strains and reagents

*Chaetomium globosum* strain W7 (CGMCC 3.14974), *Agrobacterium tumefaciens* GV3101 and pCMBIA1301-Cger plasmid were from the Microbial Genetic Engineering Lab of Harbin Institute of Technology. Restriction endonucleases and T4 DNA ligase, were purchased from Thermo Fisher Scientific. Taq DNA polymerase, hygromycin B, pMD18-T simple vector system, and *Escherichia coli* TOP10 were purchased from Takara. Standard chaetoglobosin A, pure organic reagent for chromatography were purchased from Sigma-Aldrich. Other antibiotics and acetosyringone purchased from Aladdin. Primers (Table 1) were designed using the software Primer 5.0, and synthesized by GenScript.

| Primer name | Sequence | Annealing Temperature |
|-------------|----------|-----------------------|
| iPCR1+      | 5′-TCACCTCATTAGGCCACCCC-3′  | 50 °C                |
| iPCR2-      | 5′-TGTATGGAGCAGCAGCAG-3′    | 50 °C                |
| iPCR3+      | 5′-ATCCTGTGTCGGGGTCTTG-3′   | 50 °C                |
| iPCR4-      | 5′-TCAGCGCGAAATCTCCCTC-3′   | 50 °C                |
| Cghd-F      | 5′-ATGCCTAAAACGACTGACCCC-3′ | 54 °C                |
| Cghd-R      | 5′-TTATGCCATCGCCTCATCGTC-3′ | 54 °C                |

2.2. Construction of vector

The pCMBIA1301-Cger plasmid was a vector modified from commercial pCAMBIA1301 plasmid, and the original *Gus* gene was replaced by a Cger gene (NCBI, gene No. CHGG_01240) by using Sal I and Eco91I. The pCMBIA1301-Cger plasmid was digested with Sal I, and the larger fragment was recovered. The self-connecting vector was transferred to *E. coli* Top10 for amplification, and renamed as pCMBIA1301-random. The modified plasmids were subsequently transferred into *C. globosum* by agrobacterium-mediated transformation. The processes of transformation were executed according to a previous study [10,11]. All transformants germinated from plates hygromycin B (300mg/L) were selected for further verification.

2.3. Screening of randomly inserted transformants

All the germinated transformants were transferred to the resistant plates with 500mg/L of hygromycin B. Then the transformants germinated once again were inoculated into PDB medium and cultured at 28 °C, 180 rpm for 15 d. The productions of chaetoglobosin A in all liquid mediums were detected. Detection of chaetoglobosin A was performed using an HPLC system according to a previous study [12].

2.4. Determination of T-dna insertion location

The whole genome of transformants RM48, which was chosen for the subject in further study, was extracted, and Sal I was used for single enzymatic digestion. The concentrations of purified products were determined and diluted to about 2 ng/μL. T4-DNA ligase and buffer were added to the DNA solution for overnight connection at 16 °C.

DNA purification kit was used to purify and concentrate the connected DNA. The recovery solutions were used as the templates, and the mixed DNA were amplified by PCR using primers iPCR1+/2- or iPCR3+/4-, respectively. All bands larger than 290 bp were recycled separately and sequenced. All sequences were compared with the T-DNA on pCMBIA1301-random, and the one matched with the downstream sequence of the primers iPCR1+/2- or iPCR3+/4- was able to contain...
the sequence of the potential gene inserted by T-DNA. A pair of primers Cghd-F/R was designed according to its approximate sequence (NCBI, gene No. CHGG_01271), and used to verify the correctness. Subsequently, the sequence of proteins expressed by this disrupted gene was also obtained, following total RNA extraction and reverse transcription PCR. The Neighbour-joining phylogenetic tree of similar proteins was constructed by using the software MEGA 5.0[13].

3. Results and discussion

3.1. Construction of randomly inserted plasmids

The original pCAMBIA1301 plasmid contained a *Gus* gene, which was replaced by a *Cger* gene in previous study (data not shown). The T-DNA fragment of pCAMBIA1301-Cger plasmid was too long (about 5,400 bp), reducing the chance of successful transformation. A restriction enzyme *Sal* I was chosen to shorten the plasmid (Figure 1A) based on the following reasons: large amounts in the genome (at least 13645 in genome, approximate 1/2500 bp, data from NCBI), no influence on hygromycin gene, and ability to cut off a fragment of 1809 bp.

According to the sequence of T-DNA from pCAMBIA1301-Random, four inverse PCR primers (iPCR1+, iPCR2-, iPCR3+, iPCR4-) were designed, which located in the upstream and downstream of *Sal* I, respectively. The four primers were used to detect the position of exogenous DNA fragments in the *C. globosum* genome after random insertion. The primers iPCR1+/4- can also play a role of validation in plasmid construction and transformation.

As shown in Figure 1B, the primers iPCR1+/2- can amplify the DNA at the upstream of T-DNA insertion site, while iPCR3+/4- amplify that at the downstream. Classical inverse PCR designed inverse primers directly at both the ends of the known fragment, and the PCR was performed on the circular DNA containing the entire inserted fragment, which needed a high requirement for the restriction site (the same restriction site should exist not far from the unknown fragment at both ends, and cannot be found in known DNA sequence) [14]. Therefore, it was usually necessary to change a variety of endonuclease for repeated attempts. In this modified method, a common restriction site was directly introduced into the known sequence, and the results can be detected only if one destination existed at upstream or downstream of the insertion position, which was conducive to improve the success rate of the experiment.

**Figure 1.** Plasmid pCAMBIA1301-Random and its detection process. A, construction of random mutation vector, B, process diagram of inverse PCR
3.2. Screening of transformants

T-DNA fragments were inserted into the fungal genome by the infection of \textit{A. tumefaciens}. After two screening by hygromycin B resistant plate, the final conversion rate was about 9.1/10^5 spores, and 137 resistant transformants were obtained, which were renamed as RM1 to RM137.

![Figure 2. Relative production of chaetoglobosin A in random mutation transformants.](image)

An equal number of spores from 137 transformants were inoculated into the fermentation medium, and cultured under the same conditions. The production of chaetoglobosin A from every transformant was detected, and compared with that from the wild type (set as 100%).

As shown in figure 2, the yields of chaetoglobosin A decreased for 83.2% of transformants. 16.1% of transformants had no significant change in the yields. Only one strain, named as RM48, showed a significant increase in yield to 114.8%. Therefore, RM48 was selected as the further study subject.

3.3. Determination of insertion location and function analysis of gene

The RM48 genome was extracted, digested by single enzyme and linked for PCR. No objective product was found by using primers 1+ and 2-, but in the products of PCR using primers 3+ and 4-, a DNA fragment with a size of about 1100 bp contained the known T-DNA sequence. After removing the known fragment, the rest 800 bp DNA were used for sequence alignment on NCBI. Results showed that the unknown fragment had 93% coverage rate and 99% similarity with a gene numbered CHGG01271 on the whole genome of \textit{C. globosum} CBS148.51. The following PCR by using primers Cghd-F/R designed according to gene CHGG01271 confirmed above results, and the T-DNA interrupted the gene which was similarity to CHGG01271. Real-time quantitative PCR was performed using the same concentrations of pCAMBIA1301-Random and RM48 genomes as templates, and the Ct values obtained were very close, proving that the insertion of T-DNA was a single copy. The increase in production came form the interruption of the newly discovered regulatory gene.

Phylogenetic trees used to predict gene function was drawn based on several sequences: the unknown protein, proteins with homology to unknown protein, and other proteins on the NCBI. As shown in Figure 3, the unknown protein shared high identity with the Arginase HDAC super family, and its function may be the hydrolysis of amide bonds that caused histone proteins to deacetylate and bind tightly to DNA, resulting in reduced transcriptional activity. Arginase HDAC super family had four types, and the structure of unknown protein was most close to the type I, so its gene was named as \textit{Chgd} I. The insertion of T-DNA was the 1019th nucleic acid, and the corresponding destroyed amino acid sequence was the 319th position. The destroyed amino acid just located in the HDAC conservative region (22-422), indicating that the difference between RM48 and the wild-type was caused by the destruction of the function of histone deacetylase gene \textit{Cghd} I.
4. Discussion

Due to the global regulatory effect of the genome on the synthesis of secondary metabolites, the synthesis of chaetoglobosin A will be inhibited when its extracellular accumulation reached a high concentration, leading to the failure of increase by the yields. In order to increase the final yield of active products, upstream regulatory genes must be manipulated, including two aspects: one is to overexpress the positive regulatory genes, the other is to knock out or interrupt the negative regulatory genes. According to the product repression mechanism, the destruction of the negative regulatory genes may promote the synthesis of products. Relatively few genes were known to have regulatory functions on fungal polyketide synthesis. The results showed that overexpression of the positive regulatory genes (NPG1, Lae A) could advance the synthesis time of products and may be accompanied by the generation of new active substances\(^8,9\). Lee’s research provides another way of thinking. They found the lack of histone deacetylase gene \(hda\) may increase the production of secondary metabolites\(^7\). Hu from Nankai university introduced the attempt into \(C. globosum\), but the results were not ideal\(^15\).

More negative regulation genes need to be found for the increase of the production. Under this situation, we established a detectable random insertability mutation technique based on inverse PCR and agrobacterium-mediated fungal transformation. Shorter T-DNA fragment was beneficial to improve the conversion rate, and the preservation of a high frequency cleavage site in the \(C. globosum\) genome and the provision of a high frequency cleavage site in the \(C. globosum\) genome on T-DNA and two pairs of inverse primers were beneficial to improve the success rate of validation. According to the yield change, we successfully got a strain with 14.8% of production increase and found a new gene with negative regulation on chaetoglobosin A synthesis.

The results showed that although \(Cghd\) I (similar to CHGG01271) and \(hda\) A (similar to CHGG07978) belonged to the same superfamily, they were only 28% homologous. The destruction of \(Cghd\) I and \(hda\) A had significant differences, illustrating that the former may be more critical in synthetic regulation of chaetoglobosin A.

The method used in this study relies heavily on whole-genome sequencing and needs to be further improved. However, the discovery of negative regulatory gene \(Cghd\) I indicates that this method has potential application value in finding regulatory genes.
5. Acknowledgments
This work was financially supported by the National High Technology Research and Development Program (2011AA10A205), "Twelfth Five-Year Plan" National Science and Technology Program on Rural Area (2014BAL02B00), and Doctoral Research Funding of Huainan Normal University (BSQDJ2018).

References
[1] Scherlach K, Boettger D, Remme N, et al. The Chemistry and Biology of Cytochalasans[J]. Natural Product Reports, 2010, 27(6): 869-886.
[2] Setsuko S, Kunitoshi Y, Shinsaku N, et al. Structures of Chaetoglobosin A and B, Cytotoxic Metabolites of Chaetomium globosum[J]. Tetrahedron Letters, 1973, 14(23): 2109-2112.
[3] Zhang Q, Li H Q, Zong S C, et al. Chemical and Bioactive Diversities of the Genus Chaetomium Secondary Metabolites[J]. Mini-Reviews in Medicinal Chemistry, 2012, 12(2): 127-148.
[4] Smith S, Tsai S C. The Type I Fatty Acid and Polyketide Synthases: a Tale of Two Megasynteshes[J]. Natural Product Reports, 2007, 24(5): 1041-1072.
[5] Choi S S, Nah H J, Pyeon H R, et al. Biosynthesis, Regulation, and Engineering of a Linear Polyketide Tautomycetin: a Novel Immunosuppressant in Streptomyces sp. CK4412[J]. Journal of Industrial Microbiology & Biotechnology, 2017, 44(4-5): 555-561.
[6] Kurosaki F, Togashi K, Arisawa M. Regulation of Catalytic Activity of a Multifunctional Polyketide Biosynthetic Enzyme, 6-hydroxymellitin synthase, by Interaction between NADPH and Phenylglyoxal-sensitive Amino Acid Residue at the Reaction Center[J]. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology, 2001, 1549(1): 51-60.
[7] Lee I, Oh J H, Shwab E K, et al. HdaA, a Class 2 Histone Deacetylase of Aspergillus fumigatus, Affects Germination and Secondary Metabolite Production[J]. Fungal Genetics & Biology, 2009, 46(10): 782-790.
[8] Hong J, Liu G L, Zhe C, et al. Both a PKS and a PPTase are Involved in Melanin Biosynthesis and Regulation of Aureobasidium melanogenum XJ5-1 Isolated from the Taklimakan desert[J]. Gene, 2016, 6028-15.
[9] Tao J, Wang M, Li L, et al. Overexpression of the Global Regulator LaeA in Chaetomium globosum Leads to the Biosynthesis of Chaetoglobosin Z[J]. Journal of Natural Products, 2016, 79(10): 2487.
[10] Gao X X, Yang Q. Agrobacterium tumefaciens-mediated Transformation of Chaetomium globosum and Its T-DNA Insertional Mutagenesis[J]. Acta Microbiologica Sinica, 2005, 45(1):129.
[11] Mullins E D, Chen X, Romaine P, et al. Agrobacterium-Mediated Transformation of Fusarium oxysporum: An Efficient Tool for Insertional Mutagenesis and Gene Transfer[J]. Phytopathology, 2001, 91(2):173-180.
[12] Jiang C, Song J, Zhang J, et al. New Production Process of the Antifungal Chaetoglobosin A Using Cornstalks[J]. Brazilian Journal of Microbiology, 2017:S1517838216300132.
[13] Tamura, K., Peterson, D., Peterson, N., et al. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods[J]. Mol. Biol. Evol. 2011, 28(10), 2731-2739.
[14] Ferrara F, Bradbury A R M, Sara D’Angelo. Primer Design and Inverse PCR on Yeast Display Antibody Selection Outputs[J]. Methods in Molecular Biology, 2018:35-45.
[15] Hu Y. Molecular Basis of the Secondary Metabolism in Chaetomium globosum[D]. Nankai University, 2013.