Design of tandem genes cluster and expression vector for biosynthesis of soybean isoflavones

Xia Xun-Li, Yang Guang-Xiao and He Guang-Yuan

China-UK HUST-RRes Genetics Engineering and Genomics Joint Laboratory,
Huazhong University of Science and Technology, 430074, Wuhan, China

ABSTRACT

A tandem gene cluster *CHS-CHI-IFS* (*rIFS*) for secondary metabolites of plant isoflavones was constructed by using the chalcone synthase (CHS), chalcone isomerase (CHI), and isoflavone synthase (IFS) (GenBank accession numbers EU526827, EU526829, EU526830) in a single recombination event with the pET22b vector. The resulting expression vector pET-*rIFS* was heterogeneously expressed. The highlights of the vector include ease of handling, high efficiency and universal application among diverse plant species. To the best of our knowledge, this is the first attempt at developing a novel method of constructing tandem gene cluster for future research involving secondary metabolism of isoflavones and isoflavones engineering. [*Physiol. Mol. Biol. Plants* 2009; 15(1) : 99-102] E-mail : hegy@mail.hust.edu.cn

Key words : Isoflavones biosynthesis, Novel method, Secondary metabolism, Tandem gene cluster

INTRODUCTION

In plants, secondary metabolites are involved in diverse processes like detoxification, defense responses and cell-cell communication and signaling. Secondary metabolites and their respective biosynthetic pathways are regulated by a series of genes that encode enzymes in each step of the pathway. Naturally, these are likely to have potential for secondary metabolite engineering.

Among secondary metabolites, plant isoflavones produced from *Papilionideae* Giseke of *Leguminosae* and *Euphorbiaceae*, *Ochnaceae*, *Rosaceae*, *Moraceae*, *Celastraceae*, *Amaranthaceae*, *Iridaceae*, etc. (Dewick, 1994; Harbon, 1982) play a vital role in plant physiological functions and human health. Isoflavone biosynthesis is catalyzed by Chalcone synthase (CHS), Chalcone isomerase (CHI) and Isoflavone isomerase (IFS) in a step-wise manner from coumaroyl-CoA and malonyl-CoA (Figure 1) (Brenda *et al.*, 2001; Liu *et al.*, 2002; Yu *et al.*, 2000). *CHS*, *CHI* and *IFS* encode proteins weighing 42.7 KDa, 24.9 KDa, 57.4 KDa containing 389, 226 and 521 amino acids respectively.

In the current study, we report a universally effective, novel method of constructing tandem gene cluster for secondary metabolites engineering taking plant isoflavones biosynthesis as an example. Clones were generated from the cDNAs of CHS, CHI, and IFS from soybean (*Glycine max*), and assembled into tandem gene cluster *CHS-CHI-IFS* (Designated as *rIFSs*) along with the rebuilt vector pET22b. The gene cluster was verified through the construction of an expression vector pET-*rIFS*.

METHODS AND MATERIALS

Seeds of soybean (*Glycine max*) were cultivated on wet filter paper in our laboratory. Seven days after budding, total RNA was isolated from the hypocotyls and cDNA was generated from the same. The following reagents and kits were used: Klenow fragment, T4 DNA Polymerase I, dNTP Mixture, Pyrobest DNA Polymerase, *Bam*H I, *Not* I, DNA Ligation Kit (TaKaRa); MMLV Reverse Transcriptase, RNase-H, RNase Inhibitor, dTTP, dGTP, dCTP (TOYOBO); ATP, RNA plant extraction reagent, DNA Marker 660, Protein Marker MP102, DNA Purification Kit (TIANGEN); Primers synthesis and sequencing were performed by AuGCT Biotechnology (Beijing) Co., Ltd.

Correspondence and reprint requests : He Guang-Yuan
Phone : +86-027-87792271; Fax : +86-027-87792272

Fig. 1. Plant isoflavones biosynthesis pathway

Construction of tandem gene cluster

Restriction endonuclease site of *Bam*H I of pET22b was selected and digested by *Bam*H I to generate 5' sticky ends of CTAG, which were subsequently converted to TAG using the dGTP and Klenow Fragment kit. The pET22b vector was subsequently digested by *Not* before being finally ligated with *CHS*, *CHI*, and *IFS* into the "rebuilt vector" pET22b in an orientation of *Bam*H-*CHS*-*CHI*-*IFS*-*Not*. The *CHS*, *CHI*, and *IFS* genes were cloned earlier using specially designed primers and adaptors (Figure 2).

Fig. 2. Schematic of tandem genes cluster, pET-rIFS and *CHS*, *CHI*, *IFS*'s adaptors designed

Primers design and synthesis

Primers were designed on the basis of cDNA sequences for *CHS*, *CHI*, and *IFS* (GenBank accession numbers : AY237728, AY595419 and AF195798 respectively). The actual primer sequences are as follows:

Primers of *CHS*:

chs-f: 5'-ATCGATGGTGAGCGTAGCTGAGATCCG-3';
chs-r: 5'-ACTGTCAGATGGCCACACTGCGCAG-3', the anterior of their 5' ends "ATCG" and "ACTG" were adaptors.

Primers of *CHI*:

chi-f: 5'-AGTCATGGCCACACCAGCATCCATC-3';
chi-r: 5'-AAACTCAGTTTTCAATGTTGGGATTGGC-3', the anterior of their 5' ends "AGTC" and "AAAC" were adaptors.

Primers of *IFS*:

ifs-f: 5'-TTTATGTTGCTGGAAGTTGCACTT-3'; ifs-r: 5'-GGCCATTAAGAAAGGAGTTTATAGATGCAACG-3', the anterior of their 5' ends "TTT" and "GGCCA" were adaptors.

PCR reaction conditions

For amplification of *CHS*, PCR was performed with chs-f and chs-r primers (10 μM each) 2 μl each, Reaction buffer (10 x) 5 μl, dNTP mixture (2.5 mM each) 4 μl, Pyrobest DNA polymerase (2.5U/μl) 1 μl, cDNA template 1 μl in a final volume of 50 μl. The PCR was performed at 94 °C for 3 min followed by 35 cycles of (94 °C for 45 s, 59 °C for 45 s, 72 °C for 1.5 min) and final extension at 72 °C for 15 min.

For amplification of *CHI*, PCR was performed with chi-f and chi-r primers (10 μM each) 2 μl each, Reaction buffer (10 x) 5 μl, dNTP mixture (2.5 mM each) 4 μl, Pyrobest DNA polymerase (2.5 U/μl) 1 μl, cDNA template 1 μl in a final volume of 50 μl. The PCR was performed at 94 °C for 3 min followed by 35 cycles of (94 °C for 45 s, 60 °C for 45 s, 72 °C for 1 min) and final extension at 72 °C for 15 min.

For amplification of *IFS*, PCR was performed with ifs-f and ifs-r primers (10 μM each) 2 μl each, Reaction buffer (10 x) 5 μl, dNTP mixture (2.5 mM each) 4 μl, Pyrobest DNA polymerase (2.5U/μl) 1 μl, cDNA template 1 μl in a final volume of 50 μl. The PCR was performed at 94 °C for 3 min followed by 35 cycles of (94 °C for 45 s, 52 °C for 45 s, 72 °C for 2 min) and final extension at 72 °C for 15 min.

PCR products were sequenced by AuGCT Biotechnology (Beijing) Co., Ltd.

2.5 CHS, CHI, and IFS 5' sticky ends created

The 5' sticky ends were created with the respective cDNA fragments in the following reaction mix 34.2 μ l, dCTP (4 mM) 5 μ l, BSA buffer 5 μ l, T4 DNA Polymerase buffer 5 μ l, T4 DNA Polymerase 0.8 μ l in a final volume of 50 μ l at 12 °C for 30 min, which were subsequently purified with a DNA purification kit.

Phosphorylation of 5' sticky ends of CHS, CHI, and IFS

The 5' sticky ends of *CHS*, *CHI*, and *IFS* were phosphorylated with cDNA fragments of *chs* (or *chi*, or *ifs* respectively) in the following reaction: 12 μ l, Buffer (10x) 2 μ l, ATP (10 mM) 5 μ l, T4 Polynucleotide Kinase 1 μ l in a final volume of 50 μ l at 37 °C for 60 min. These were subsequently purified using the DNA Purification kit.

Ligation of Rebuilt vector pET22b and CHS, CHI, IFS and transformation

The 5'-phosphorylated fragments of *CHS*, *CHI*, and *IFS* with sticky ends were assembled and ligated into pET22b between *Bam*H I and *Not* I sites in an orientation of *Bam*H-*CHS*-*CHI*-*IFS*-*Not* I. The reaction set up comprised of 1 μ l each of cDNA fragments of *CHS*, *CHI*, and *IFS*, 1 μ l rebuilt pET22b vector, 4 μ l DNA Ligation Kit in a final volume of 8 μ l at 16 °C for 2 hrs after which they were transformed into 100 μ l competent cell of DH5 α for positive clones screening.

Fig. 3. Electrophoresis of PCR of cDNA of *CHS*, *CHI*, and *IFS* (M=DNA Marker, the sizes are 4500 bp, 3000 bp, 2000 bp, 1200 bp, 800 bp, 500 bp, and 200 bp from the top down respectively; A: *CHS* = 1170 bp, B: *CHI* = 681 bp, C: *IFS* = 1566 bp)

Fig. 4. Electrophoresis of PCR of *CHS*, *CHI* and *IFS* of pET-*rIFS* (M = DNA Marker, the sizes are 4500 bp, 3000 bp, 2000 bp, 1200 bp, 800 bp, 500 bp, and 200 bp from the top down respectively; D: *CHS* = 1170 bp, E: *CHI* = 681 bp, F: *IFS* = 1566 bp).

RESULTS

PCR confirmation of cDNA of CHS, CHI, and IFS

The PCR analyses for the genes were in agreement with the GenBank data for the respective genes, which were further, confirmed using BLAST and accorded over 99 % similarity with the query sequences (Figure 3).

PCR confirmation of CHS, CHI, and IFS of pET22-rIFS

PCR analyses were carried out to test the positive clone for selecting the successful expression vectors (Figure 4).

Expression studies of pET-rIFS

pET-*rIFS* vector plasmid containing Ampicillin as resistance marker was transformed into DE3 strain and

Fig. 5. SDS-PAGE analysis of pET-*rIFS* (M = Protein Marker MP102; 1,2,3,4 are samples of pET-*rIFS*, CHS = 42.7 KDa, CHI = 24.9 KDa, IFS = 57.4 KDa)

grown in LB medium (200 µg/mL Ampicillin) with IPTG (0.3 mmol/L) as an inducer. SDS-PAGE analyses of the expressed proteins confirmed the respective molecular weights of the proteins (Figure 5).

CONCLUSION AND DISCUSSION

Plant secondary metabolic reactions are catalyzed by a series of enzymes, making single gene or enzyme based approaches to secondary metabolites engineering for recombinant crops with increased secondary products redundant. Cloning and assembling multiple genes of secondary metabolite synthesis pathway into a single tandem gene cluster in an orientation as reported above offers alternatives for secondary metabolites engineering with tandem gene clusters. This approach is more reliable and effective than constructing multi-gene expression vectors by recombining genes.

Efforts to construct recombinant vectors are most often compounded due to lack of compatible restriction sites and options for inserting and screening the desired fragments. In our novel approach, we did not use any restriction endonuclease to digest the *CHS*, *CHI*, and *IFS* to get the needed sticky ends. Instead, we obtained sticky ends through a PCR based approach, using specific adaptors incorporated into gene specific primers.

This PCR based approach has more far-ranging applications for proper restriction endonuclease site(s) selection or making point nucleotide mutation to partially

smoothen the 5' sticky ends in the presence of dNTPs. However, care should be taken to ensure that the desired combination of restriction endonucleases generates sticky ends with not less than four nucleotides and at least one of the two has to be a non palindromic in nature. This is to ensure high ligation efficiency since sticky ends with less than three nucleotides drastically reduce the same.

In conclusion, the recombinant expression vector pET22b has been shown to be a highly efficient and reliable system for recombinant studies involving tandem gene clusters and has a very high degree of applicability for secondary metabolite engineering.

REFERENCES

- Brenda WS (2001). Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol.*, 126: 485-493.
- Dewick PB (1994). The Flavonoids. London Chapman & Hall.
- Harbon JB (1982). The Flavonoids: Advances in Research London. Chapman & Hall New York. 537-559.
- Liu CJ, Jack WB, Christopher LS and Dixon RA (2002). Bottlenecks for metabolic engineering of isoflavone glycoconjugates in *Arabidopsis*. *PNAS*, 99(22): 14578-14583.
- Yu O, Jung W, Shi J, Robert AC, Gary MF, Brian MG and Joan TO (2000). Production of the isoflavones genistein and daidzein in non-legume dicot and monocot tissues. *Plant Physiol.*, 124: 781-793.