

ORIGINAL ARTICLE

Open Access



# Rapid monitoring of the target protein expression with a fluorescent signal based on a dicistronic construct in *Escherichia coli*

Chang-Ye Hui<sup>\*</sup> , Yan Guo, Wen Zhang and Xian-Qing Huang

## Abstract

Real-time quantification of recombinant proteins is important in studies on fermentation engineering, cell engineering, etc. Measurement of the expression level of heterologous proteins in bacterial fermentation broth has traditionally relied on time-consuming and labor-intensive procedures, such as polyacrylamide gel electrophoresis, immunoblot analysis, and biological activity assays. We describe a simple, fast, and high sensitive assay for detecting heterologous proteins production in bacteria either at the overall level (fluorescence spectrophotometry) or at the individual level (fluorescence microscopic image) in this study. Based on a dicistronic model, the translation of target gene in the upstream open reading frame (ORF) was coupled with the synthesis of the mCherry reporter in the downstream ORF in *E. coli* cells, and subsequently this demonstrated a positive correlation between the expression of target gene and mCherry. Although a time lag exists between the expression of target protein and mCherry reporter, the method described here allows facile monitoring of dynamic changes in target protein expression, relying on indirect determination of the fluorescence intensity of mCherry during fermentation in real-time models. Additionally, the performance of a single bacterial cell factory could be checked under the fluorescence microscope field.

**Keywords:** Recombinant protein expression, Dicistron, Fluorescent signal, Real-time, Quantification

## Introduction

It is a long time since *Escherichia coli* (*E. coli*) was developed as an ideal microbial cell factory for the production of recombinant proteins. The application of highly active bacteriophage RNA polymerase, especially T7 RNA polymerase, made the expression system very powerful for overproduction of recombinant proteins (Peranen et al. 1996). An *E. coli*-based expression system is certainly used as the preferred system, and numerous heterologous proteins of different origins have been successfully expressed in recombinant *E. coli* (Hui et al. 2018b; Mahalik et al. 2014).

The expression level of recombinant protein can be affected by many factors, including genetic elements, secondary structure of mRNA, codon preference, culture

conditions, and more (Mahalik et al. 2014). Measurement of the expression level of recombinant proteins in fermentation broth has generally relied on time-consuming and labor-intensive routine assays such as PAGE, western blot, ELISA, and biological activity assay. Currently, there are few approaches to precisely detect the expression level of the target protein, not to mention the technique for real-time monitoring of the translation level of the recombinant protein.

Operons are multigene transcriptional units which occur primarily in prokaryotes but also in some eukaryotes (Blumenthal 2004; Jacob et al. 2005). A typical bacterial operon was defined as several structural genes arranged in tandem that are transcribed in a single polycistronic mRNA under the control of the same promoter (Jacob et al. 2005). However, naturally occurring genetic regulatory elements hardly meet the requirements for biotechnological applications. Progress in the development of artificial hybrid operons promises to greatly

\*Correspondence: hcy\_sypu@hotmail.com  
Department of Pathology & Toxicology, Shenzhen Prevention and Treatment Center for Occupational Disease, Shenzhen 518020, People's Republic of China

broaden the use of *E. coli* as a model for the production of heterologous proteins (Yadav et al. 2014).

In this paper, we have developed a molecular device in *E. coli* using a dicistronic construct. A short ORF containing multiple cloning sites (MCS) was placed upstream of another ORF encoding red fluorescence protein mCherry as reporter. The target gene can be ligated into MCS, and both the target gene and the reporter gene are co-transcribed into one dicistronic mRNA. More importantly, the expression of the target gene would also be coupled with the expression of the reporter mCherry. Thus, the target protein production could be easily quantitated by indirect determination of the fluorescence intensity of mCherry in crude *E. coli* culture in real-time. The fluorescent signal can be noninvasively detected in living *E. coli* cells during culture at any time. This eliminates the analysis of exogenously expressed protein by routine procedures. In addition, whether the recombinant protein is expressed or not, which can be conveniently determined by fluorescence microscopic imaging with a standard fluorescence microscope. The strategy described in this study greatly simplifies the monitoring of dynamic changes in target protein expression during fermentation.

## Materials and methods

### Bacterial strains and agents

The *E. coli* strains and plasmids used in this study are shown in Table 1. *E. coli* Top10 was used for all the cloning steps, while BL21(DE3)pLysS and Rosetta(DE3)pLysS

were used for recombinant protein expression. Restriction enzymes, Pyrobest DNA polymerase, dNTP and gel extraction kits were obtained from TaKaRa (Dalian, China). Isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG), and antibiotics were purchased from Sangon Biotech (Shanghai, China). Tryptone and yeast extract were obtained from OXIOD (Basingstoke, UK). *E. coli* was cultured in Luria–Bertani (LB) broth (1% tryptone, 0.5% yeast extract and 1% NaCl) supplemented with antimicrobial agents, as necessary.

### Construction of recombinant plasmids

DNA encoding red fluorescent protein mCherry (GenBank Accession No. MH070102) was optimized based on the *E. coli* codon bias and synthesized by Sangon Biotech (Shanghai, China). The synthesized fragment (708 bp) was cloned into pUCm-T to generate pT-RFP. All primers used in this study are shown in Table 2. The gene encoding mCherry was amplified by PCR from pT-RFP using the 1mF and 1mR primers, digested with *Nde*I and *Bam*HI, and then ligated into similarly digested pET-21a to yield pET-RFP, which produced recombinant mCherry under the control of T7 lac promoter. A 0.72 kb DNA fragment containing *mcherry* was amplified from pT-RFP using the 2mF and 2mR primers, then digested and ligated into the *Hind*III-*Xho*I sites of pET-21a, forming pO-RFP. DNA fragments (about 0.75 kb) were amplified from pO-RFP using the 2m1F/2m2F/2m3F and 2mR primers, then digested and inserted in pO-RFP via *Eco*RI and *Xho*I restriction sites, generating pOI-RFP/pOII-RFP/pOIII-RFP.

**Table 1 Bacterial strains and plasmids used in this study**

	Genotype/characteristics	Source
Strains		
Top10	F <sup>-</sup> <i>mcrA</i> $\Delta$ ( <i>mrr-hsdRMS-mcrBC</i> ) $\phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ <i>lacX74</i> <i>recA1</i> <i>araD139</i> $\Delta$ ( <i>ara-leu</i> )7697 <i>galU</i> <i>galk</i> <i>rpsL</i> (Str <sup>R</sup> ) <i>endA1</i> <i>nupG</i>	Invitrogen
BL21(DE3)pLysS	F <sup>-</sup> <i>ompT</i> <i>hsdS</i> <sub>g</sub> ( <i>r</i> <sub>g</sub> <sup>-</sup> <i>m</i> <sub>g</sub> <sup>-</sup> ) <i>gal dcm</i> (DE3) pLysS (Cm <sup>R</sup> )	Novagen
Rosetta(DE3)pLysS	F <sup>-</sup> <i>ompT</i> <i>hsdS</i> <sub>g</sub> ( <i>r</i> <sub>g</sub> <sup>-</sup> <i>m</i> <sub>g</sub> <sup>-</sup> ) <i>gal dcm lacY1</i> (DE3) pLysSRARE (Cm <sup>R</sup> )	Novagen
Plasmids		
pET-21a	Amp <sup>r</sup> T7 promoter lac operator	Novagen
pUCm-T	TA cloning	Sangon
pT-RFP	pUCm-T carrying <i>mcherry</i>	This study
pET-RFP	pET-21a expressing mCherry	This study
pO-RFP	pET-21a carrying <i>mcherry</i> in ORF2	This study
pOI-RFP	pO-RFP with one rare codon in ORF1	This study
pOII-RFP	pO-RFP with two consecutive rare codons in ORF1	This study
pOIII-RFP	pO-RFP with three consecutive rare codons in ORF1	This study
pT-GFP	pUCm-T carrying <i>egfp</i>	This study
pGFP-RFP	pO-RFP carrying <i>egfp</i> in ORF1	This study
pGFPI-RFP	pGFP-RFP with one rare codon preceding <i>egfp</i>	This study
pGFPII-RFP	pGFP-RFP with two consecutive rare codons preceding <i>egfp</i>	This study

The gene of eGFP (GenBank Accession No. MH070103) was synthesized according to the codon usage bias of *E. coli*, and subcloned into pUCm-T to generate pT-GFP. A 0.73 kb DNA fragment containing *egfp* was amplified from pT-GFP using the 2eF and 2eR primers, then digested and ligated into the *Bam*HI–*Hind*III sites of pO-RFP, forming pGFP-RFP. DNA fragments (about 0.73 kb) were amplified from pGFP-RFP using the 2e1F/2e2F and 2eR primers, then digested and inserted in pGFP-RFP via *Bam*HI and *Hind*III restriction sites, generating pGFPI-RFP/pGFPII-RFP. All recombinant plasmids were sequenced (Sangon Biotech, China) to confirm accordance with the design.

### Protein expression

*Escherichia coli* expression host BL21(DE3)pLysS or Rosetta(DE3)pLysS was transformed with recombinant vectors using a CaCl<sub>2</sub>-mediated transformation method (Hui et al. 2018a). The transformed *E. coli* cells were spread on LB agar plates containing 50 µg/mL ampicillin and 34 µg/mL chloramphenicol, and then cultured overnight at 37 °C.

A single colony was used to inoculate 3 mL of LB medium in a 15 mL conical tube supplemented with the above antibiotics, and cultured overnight in a 37 °C shaking incubator at 200 rpm. For optimal protein expression experiment, the pre-cultured recombinant *E. coli* as described above was diluted to an OD<sub>600</sub> of 0.01 in 100 mL fresh LB broth containing antibiotics in 250 mL triangular flasks. The culture was grown at 37 °C for 3 h with rotation at 220 rpm, and the bacteria reached log phase with an optical density at 600 nm of 0.4–0.5. The cultures were then induced for 6 h with 0.5 mM IPTG

and incubated at 37 °C with shaking at 220 rpm. The 5 mL cultures were sampled at regular intervals after induction, and OD<sub>600</sub> was measured by using an iMark microplate reader (Bio-Rad, USA).

### Fluorescence quantitative analysis

The fluorescence intensity of eGFP or mCherry produced in *E. coli* was measured with Lumina fluorescence spectrometer (Thermo, USA). A 3-mL aliquot of *E. coli* culture or diluent was added to 1-cm low fluorescence background quartz cuvette. To test the fluorescence intensity of recombinant proteins, the excitation wavelength was set at 488 nm and the intensity of emitted fluorescence of eGFP at 507 nm (Tsien 1998) was recorded. Subsequently, the excitation wavelength was set at 587 nm and the intensity of emitted fluorescence of mCherry at 610 nm (Shaner et al. 2004) was recorded at the same time. The fluorescence intensity was normalized by dividing the fluorescence intensity at the emission wavelength of 507 nm (eGFP) or 610 nm (mCherry) by the OD<sub>600</sub> value of the same sample.

### Fluorescence microscopy

An aliquot of 50 µL of *E. coli* culture was directly mounted onto a glass slide, and air-dried. *E. coli* cells were visualized using a Nikon Eclipse Ni fluorescence microscope coupled to a Nikon DS-Ri2 digital camera (Tokyo, Japan). A FITC filter (excitation at 475–490 nm and emission at 505–535 nm) was used to detect eGFP fluorescence, and a TRITC filter (excitation at 545–565 nm and emission at 580–620 nm) was used to detect mCherry fluorescence in *E. coli* cells. Representative images were captured using NIS Elements software (Nikon, Japan).

## Results

### Construction of a series of dicistronic vectors

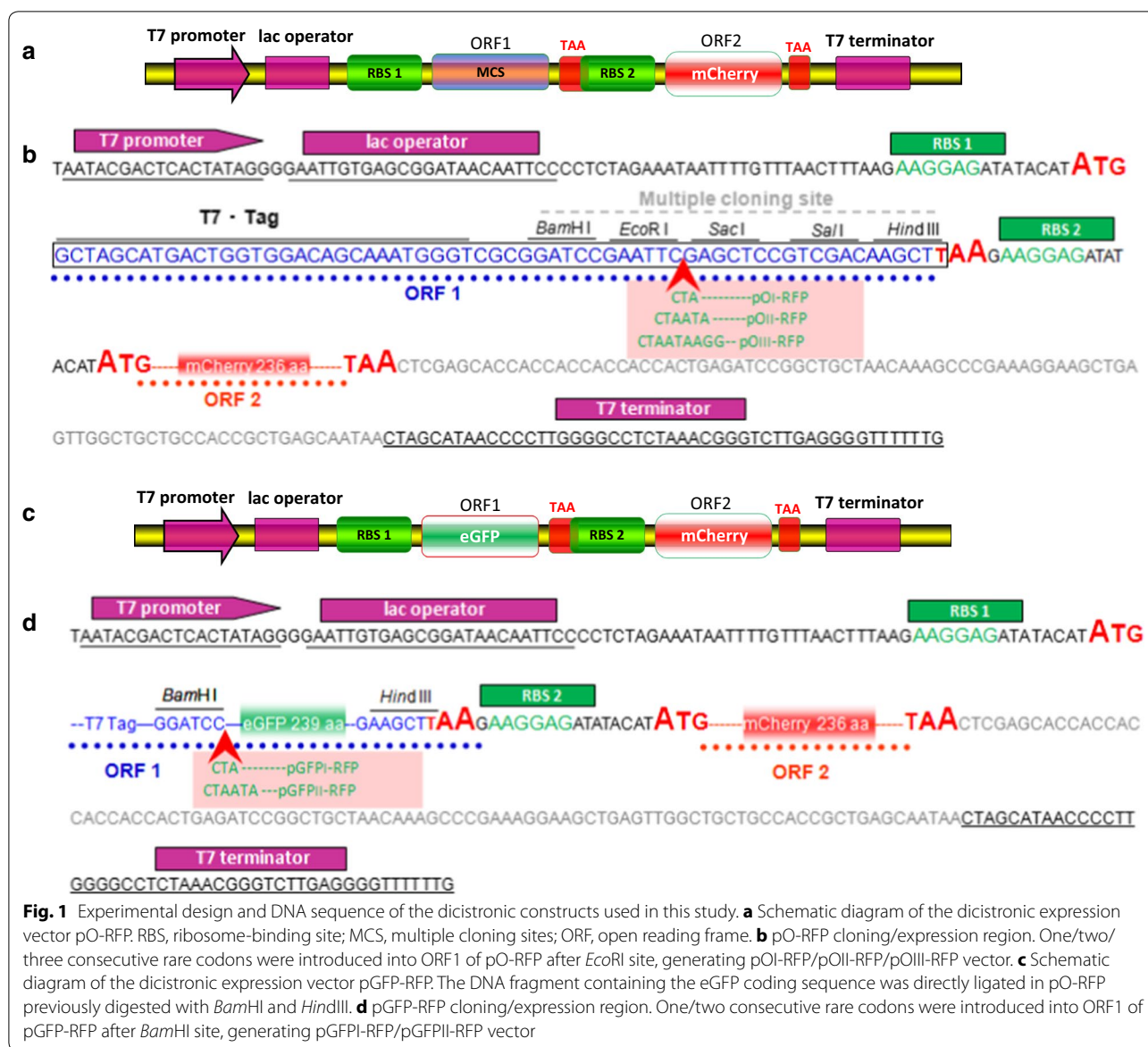
The strategy used in this study to construct the dicistronic vectors is shown in Fig. 1. The DNA fragment containing extra strong RBS, ATG initiation codon and the mCherry coding sequence was directly ligated in pET-21a previously digested with *Hind*III and *Xho*I, forming a dicistronic construct pO-RFP, which produced mCherry in the second ORF (ORF2). One/two/three consecutive rare codons were introduced following *Eco*RI site in the first ORF (ORF1) based on the PCR method, then generating pOI-RFP/pOII-RFP/pOIII-RFP, respectively.

The DNA fragment containing the eGFP coding sequence was directly inserted into ORF1 of pO-RFP to yield a dicistronic construct pGFP-RFP, which produced the target protein eGFP in ORF1 and the reporter protein mCherry in ORF2. One/two consecutive rare codons were placed upstream of *egfp* based on the PCR method, then generating pGFPI-RFP/pGFPII-RFP, respectively.

**Table 2 Primers used in this study**

Primer	Sequence 5'–3'	Restriction site
1mF	GGATCTGCATATGGTCTCTAAAGGCGAGGAAG	<i>Nde</i> I
1mR	CGGGATCCTTATTTGTACAGTTCGTCATG	<i>Bam</i> HI
2mF	CCCAAGCTTAAGAAGGAGATATACATATG GTCTCTAAAGGCG	<i>Hind</i> III
2mR	CCGCTCGAGTTATTTGTACAGTTCGTCATG	<i>Xho</i> I
2m1F	GGAATTC <b>CTA</b> GAGCTCCGTCGACAAG	<i>Eco</i> RI
2m2F	GGAATTC <b>CTAATA</b> GAGCTCCGTCGACAAG	<i>Eco</i> RI
2m3F	GGAATTC <b>CTAATAAGG</b> GAGCTCCGTCGACAAG	<i>Eco</i> RI
2eF	CGGGATCCATGGTAAGCAAGGGTGAG	<i>Bam</i> HI
2eR	CCCAAGCTCTTTGTACAGTTCATCCATG	<i>Hind</i> III
2e1F	CGGGATCCATG <b>CTA</b> GTAAGCAAGGGTGAG	<i>Bam</i> HI
2e2F	CGGGATCCATG <b>CTAATA</b> GTAAGCAAGGGTGAG	<i>Bam</i> HI

Rare codons are represented in bold, and restriction enzyme sites are indicated with underline

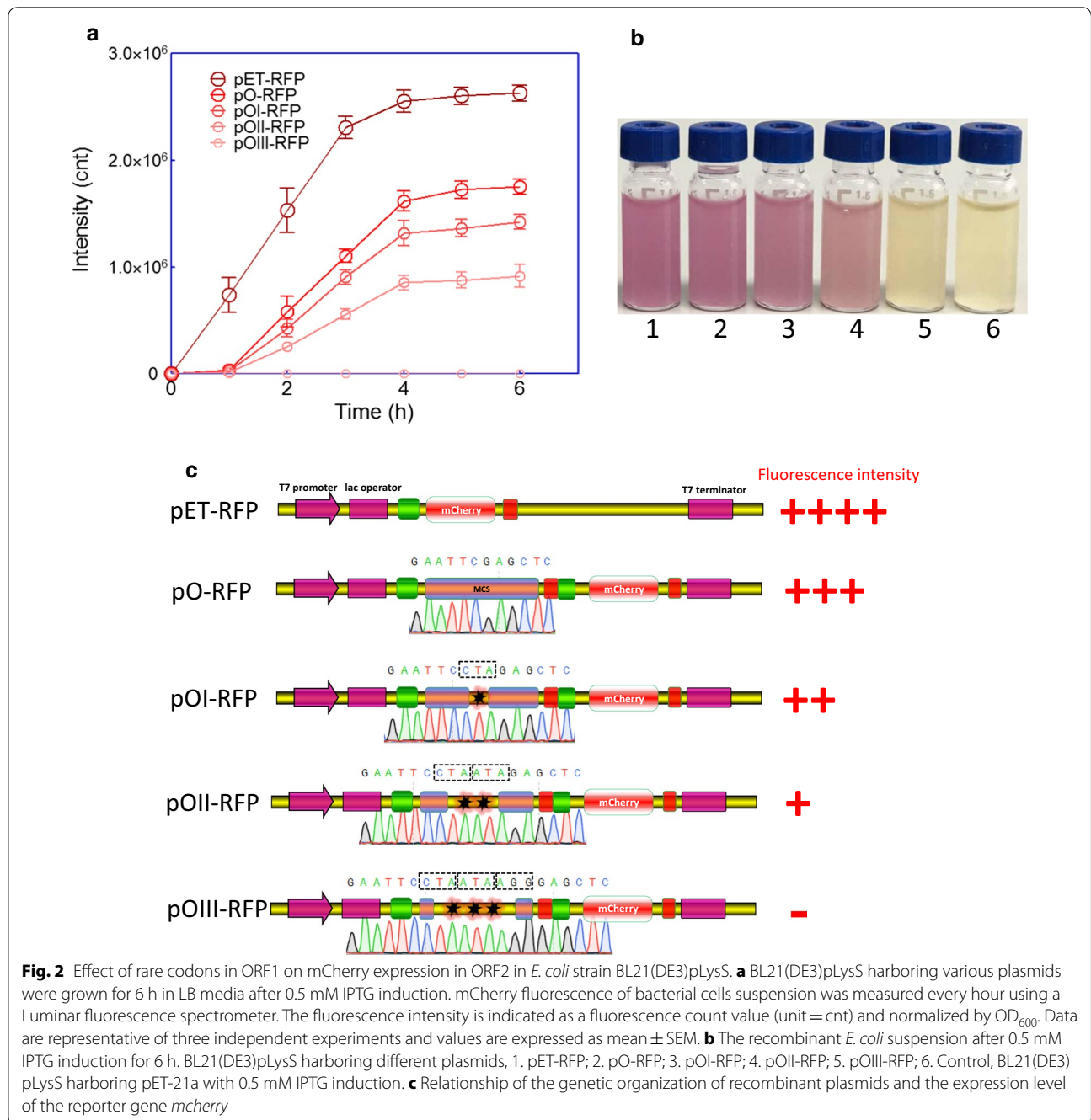


### mCherry expression in ORF2 was tuned by rare codons in ORF1

The log phase cultures of BL21(DE3)pLysS harboring pET-RFP, pO-RFP, pOI-RFP, pOII-RFP, and pOIII-RFP were induced for 6 h with 0.5 mM IPTG at 37 °C, and there was no significant difference in the slope of the growth curve in these five recombinant *E. coli*. The OD<sub>600</sub> of the cultures increased continuously since initiation of induction, and nearly reached a stable state after 4 h (data not shown).

In the dicistronic vector pO-RFP, a short ORF1 composed of T7 Tag coding sequence and MCS, encoding a 22-amino acid peptide (the amino acid sequence of encoding polypeptide is

MASMTGGQQMGRGSEFELRRQA), was placed upstream of ORF2 encoding mCherry. Compared with BL21(DE3)pLysS/pET-RFP, insertion of upstream short ORF1, resulted in significant decrease of mCherry expression (Fig. 2a). Although about 1 h mCherry expression lag resulted from the inserted short ORF1 existed, no significant improvement in mCherry expression was observed in all recombinant *E. coli* after 4 h induction (Fig. 2a). By changing the amount of rare codons inserted into ORF1, we sought to change the translation level of ORF1 and subsequently affect mCherry expression in ORF2. As expected, the introduction of rare codons in ORF1 lead to suppression of mCherry expression in ORF2, which was reflected by decreased fluorescence



intensity throughout the induction time course (Fig. 2a). mCherry overexpressed in *E. coli* made the cultures light up in red, and color difference at 6 h terminal point was significantly obvious among these groups (Fig. 2b). The number of rare codons in ORF1 was negatively correlated with the expression of mCherry in ORF2 (Fig. 2c). Interestingly, three consecutive rare codons placed in ORF1 effectively abolished the expression of mCherry.

**Tuning mCherry expression in ORF2 with rare codons in ORF1 was partially recovered by Rosetta(DE3)pLysS host**  
 The expression of gene containing rare codons can be dramatically improved when the cognate tRNAs are increased within the host (Mahalik et al. 2014). The Rosetta strains contain vector pRARE, which encodes the rare tRNA genes (Madje et al. 2012; Nguyen et al. 2011). Thus, the rare codons CUA, AUA, and AGG involved in pOI-RFP/pOII-RFP/pOIII-RFP should all

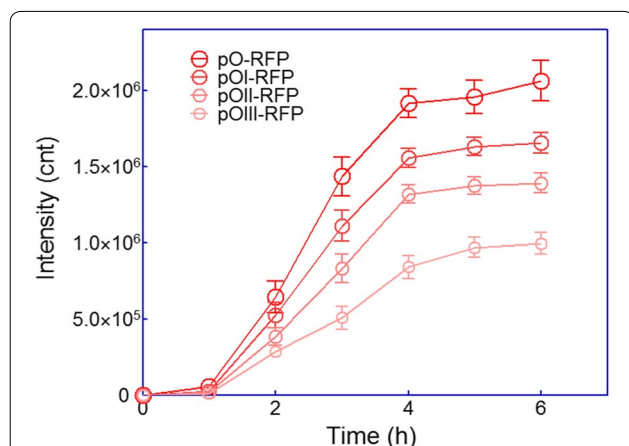


be supplemented. When the host was transferred from BL21(DE3)pLysS to Rosetta(DE3)pLysS, the expression of ORF1 containing rare codons was expected to be recovered due to the supply of rare tRNAs.

The log phase cultures of Rosetta(DE3)pLysS harboring four dicistronic plasmids, including pO-RFP, pOI-RFP, pOII-RFP, and pOIII-RFP, were induced for 6 h with 0.5 mM IPTG at 37 °C. The growth curve of recombinant Rosetta(DE3)pLysS was similar with that of recombinant BL21(DE3)pLysS, and there was no significant difference among four recombinant bacteria strains (data not shown). Although there were still significant differences in the red fluorescence intensity among four recombinant Rosetta(DE3)pLysS strains, obvious restored expression of mCherry in ORF2 was observed especially in Rosetta(DE3)pLysS/pOIII-RFP. The results are shown in Fig. 3.

#### eGFP expression in ORF1 was positively correlated with mCherry expression in ORF2

The dicistronic vector pO-RFP can be assembled by cloning target gene into the MCS located in ORF1. In order to test whether the expression of target gene in ORF1 can be indicated by the fluorescence reporter mCherry, the model protein eGFP coding sequence as the target gene was subsequently inserted into the MCS, substituting for the original short ORF1. Because there is no overlap of the excitation and emission spectra of mCherry and eGFP, it is convenient to detect the expression of mCherry and eGFP simultaneously by a fluorescence method (Shaner et al. 2004; Tsien 1998).



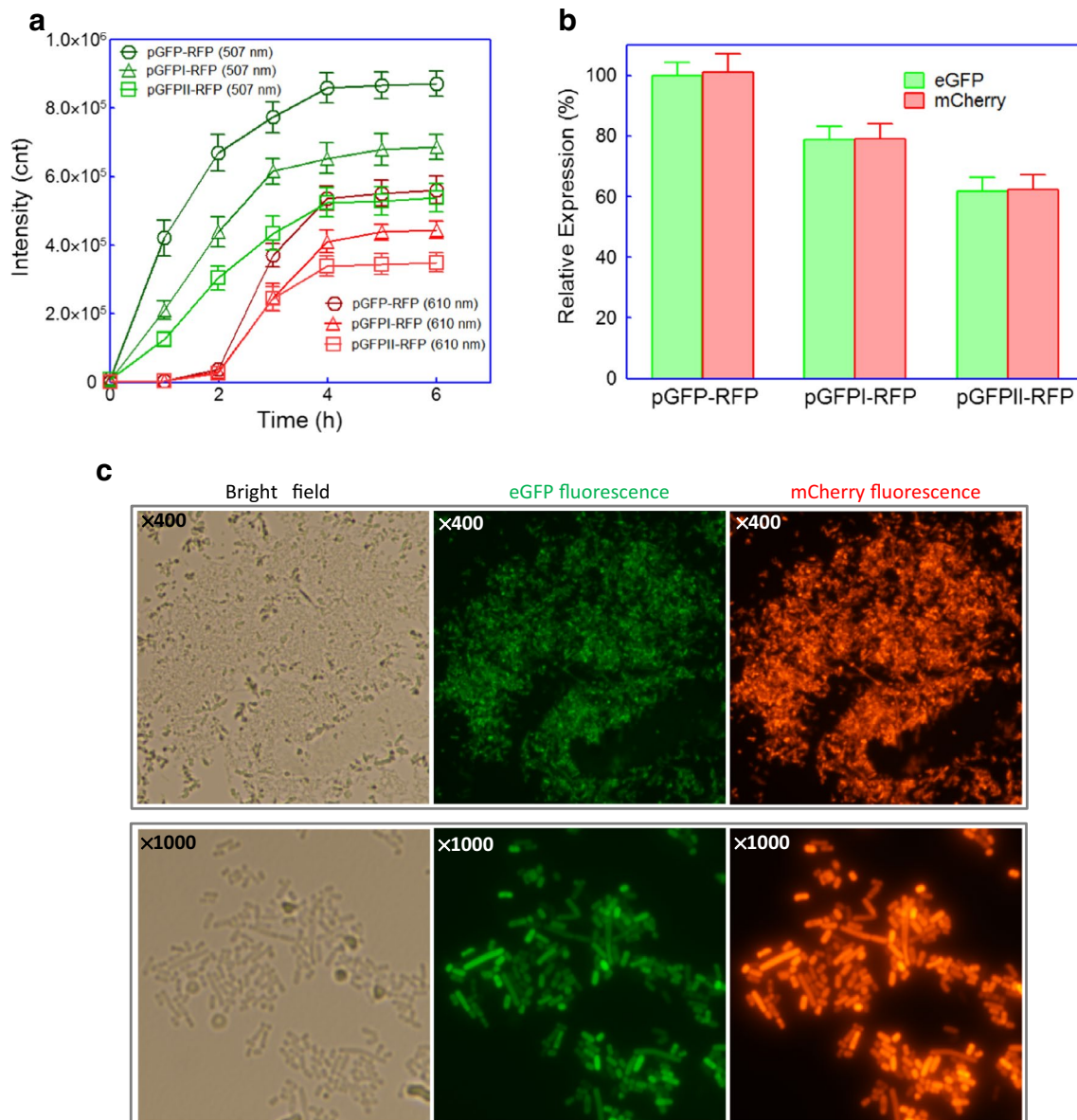
**Fig. 3** Effect of rare codons in ORF1 on mCherry expression in ORF2 in *E. coli* strain Rosetta(DE3)pLysS. Rosetta(DE3)pLysS harboring various plasmids were grown for 6 h in LB media after 0.5 mM IPTG induction. mCherry fluorescence of bacterial cells suspension was measured every hour using a Lumina fluorescence spectrometer. Fluorescence intensity values were normalized using the absorbance at 600 nm. Data are representative of three independent experiments and values are expressed as mean  $\pm$  SEM

The log phase cultures ( $OD_{600}$  = 0.4–0.5) of BL21(DE3)pLysS harboring pGFP-RFP, pGFPI-RFP, and pGFPII-RFP were induced with 0.5 mM IPTG at 37 °C, and the different kinds of dicistronic plasmids exerted no obvious effects on the growth of recombinant *E. coli* after induction (data not shown). Green fluorescence in all three recombinant bacterial cultures could be detected immediately at the early stage of induction, but there was about 1 h lag of mCherry expression after IPTG induction. No significant increase in both eGFP and mCherry expression was observed in all recombinant *E. coli* after 4 h induction (Fig. 4a). Rare codons placed near the N-terminus of the well-expressed protein eGFP were demonstrated to downregulate eGFP expression effectively, which is reflected in the eGFP fluorescence intensity (Fig. 4a). Interestingly, followed by downregulation of eGFP expression in ORF1, the decrease of mCherry expression in ORF2 was also observed. Expression level of the target protein eGFP could be reflected by red fluorescence intensity directly during induction (Fig. 4a) or at the end of induction (Fig. 4b). In addition, bacterial whole cell fluorescence imaging could be done based on both the target protein eGFP and the reporter protein mCherry. *E. coli* cells were conveniently visualized using a fluorescence microscope, and the fluorescence intensity in a single bacterial cell factory could be distinguished under 1000 times magnification (Fig. 4c).

#### Discussion

In *E. coli*, RBS and other regulatory RNA sequences are essential control elements for translation initiation (Carrier and Keasling 1999; Isaacs et al. 2004). Translation initiation, the rate-limiting step of protein expression, is determined by several molecular interactions, such as the complementarity of the RBS sequence to 16S rRNA, the spacing between RBS and the start codon, and the presence of mRNA secondary structures (Salis et al. 2009). In some cases, the distance between the promoter and structural genes in natural operons is 60 bases above. The necessary elements for translation initiation including: the standby site, 16S rRNA binding site, spacing and the start codon are all included in the intergenic regions. So each gene in the polycistronic mRNA strand can be translated independently and tunably (de Smit and van Duin 2003; Jacob et al. 2005; Kudla et al. 2009).

The underlying goal of synthetic biology is to develop concise and standardized biological parts, so no redundant sequences are retained (Shetty et al. 2008). How do we ensure that the downstream genes in the artificial operon can be translated successfully? The Shine–Dalgarno (SD) sequences are obviously necessary but not sufficient. Several effective strategies were developed to address this issue, including: a large fragment of optimal



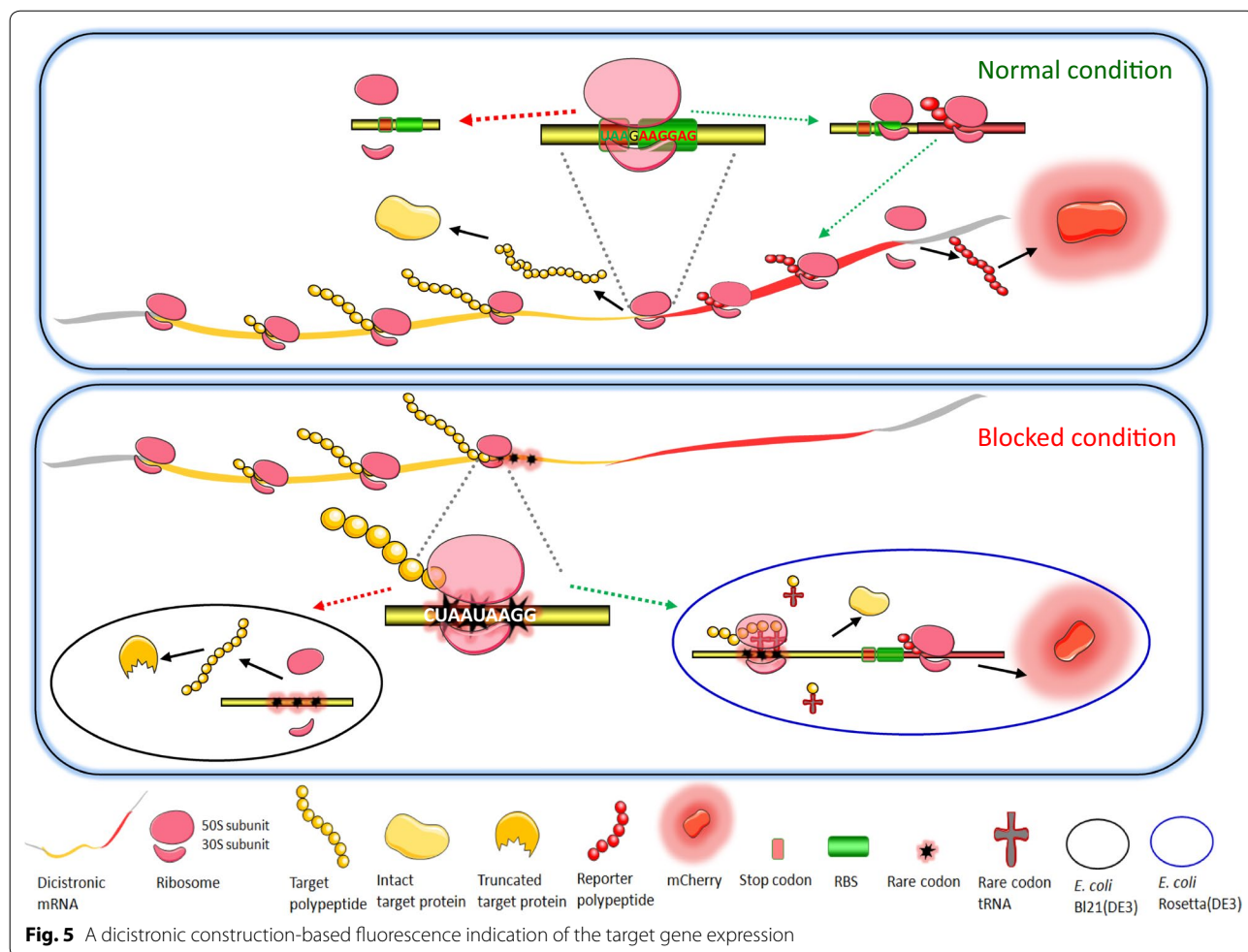
**Fig. 4** Relationship of the expression level of the target gene *egfp* in ORF1 and the reporter gene *mcherry* in ORF2. **a** Time course of the fluorescence intensity of eGFP and mCherry. BL21(DE3)pLysS harboring pGFP-RFP, pGFPI-RFP, and pGFPII-RFP were grown for 6 h in LB media after 0.5 mM IPTG induction. **b** Comparison of the co-expression levels of eGFP and mCherry in three recombinant *E. coli* at 6 h after IPTG induction. The fluorescence values of eGFP and mCherry in pGFP-RFP were defined as 100%, and the fluorescence values in pGFPI-RFP and pGFPII-RFP were normalized by dividing the actual fluorescence values by the fluorescence values of eGFP and mCherry in pGFP-RFP. Data are representative of three independent experiments and values are expressed as mean  $\pm$  SEM. **c** Fluorescence microscopic analysis of *E. coli* BL21(DE3)pLysS/pGFP-RFP after 4 h induction with 0.5 mM IPTG. *E. coli* cells were visualized using a fluorescence microscope ( $\times 400$  or  $\times 1000$  magnification). The fluorescence imaging of eGFP upon excitation at 475–490 nm and emission detection between 505 and 535 nm, and fluorescence imaging of mCherry upon excitation at 545–565 nm and emission detection between 580 and 620 nm. The same setting was applied for imaging of both eGFP and mCherry. At least three independent preparations were analyzed and representative images are shown here

or natural DNA sequence containing the elements for the translation initiation was placed in the intergenic regions (Wang et al. 2015; Yadav et al. 2014); the coding sequence or the stop codon of the upstream gene was overlapped with the SD sequence of the downstream gene (Ito and

Kurosawa 1992; Palleja et al. 2009); and the SD sequence of the downstream gene was placed immediately after the termination codon of the upstream gene (Esipov et al. 2016; Gaymard et al. 2000).

In order to achieve the co-regulation of structural genes, functionally related genes that are common in both prokaryotic and eukaryotic genome are often found in close proximity (Gasch and Eisen 2002; Palleja et al. 2009). Inspired by this phenomenon, previous studies have demonstrated that the expression of structural genes could be co-regulated in an artificial operon by assembling them in tandem in close proximity, and high expression was always achieved in the upstream gene (Cheng et al. 2016; Esipov et al. 2016; Wang et al. 2015). The distance between the stop codon of the upstream gene and the start codon of the downstream gene was demonstrated to be of key importance in translational coupling (Levin-Karp et al. 2013). These findings suggest that translation of co-regulated structural genes was initiated not only in order but also coupled together. To develop a simple approach for indicating the target protein expression of *E. coli* in real-time, we engineered a short ORF containing MCS for inserting the target gene, which was placed upstream of another ORF encoding the reporter fluorescence protein. To achieve a high-level

co-expression of reporter mCherry with the target gene, an optimized strong RBS and spacing between SD sequence and ATG initiator codon (Pfleger et al. 2006) were placed immediately after the termination codon of the first ORF. Molecular mechanism of the hypothesis involved in this study is shown in Fig. 5. It allows the two genes to be co-transcribed into one dicistronic mRNA strand under the control of T7 lac promoter, and not only independent but also coupling production of target protein and reporter. No extra or enough ribosomes might bind RBS2 directly, because not enough characteristic DNA sequences related to translation initiation were included in the short length between the target gene and the report gene. During normal conditions when the ribosome moves to the first terminal codon, nascent target polypeptide will be released. However, the ribosome tends to dissociate. Due to the next RBS immediately following the terminal codon, a few residual ribosomes that have not yet dissociated will continue to initiate the expression of the reporter. During the blocked condition, a variety of possible factors, such as consecutive





rare codons in the first ORF, may lead to failure in target gene expression. Most stalled ribosome will then be subsequently released due to tRNA non-availability, and the mCherry reporter expression will be closed. Another important mechanism involved in the coupled translation of tandem cistrons might be that highly efficient polycistronic mRNA degradation is always triggered by substrate starvation, depletion of nutritional elements, etc. (Mahalik et al. 2014). The dicistronic mRNA degradation could be conveniently monitored using real time PCR during blocked conditions.

Rare codons generally influence the rate and efficiency of translation, due to rarity of their cognate tRNAs (Kane 1995; Parmley and Huynen 2009; Rosenberg et al. 1993). Protein expression could be regulated precisely by changing the combination of genetic codes in both mammalian cell and engineered bacterial cells (Brule and Grayhack 2017; Wang et al. 2016). The rare codons involved in this study (CTA, ATA and AGG) were used at a frequency of <0.5% in *E. coli*, and previous studies have demonstrated that the above rare codon alone could cause translational problems in *E. coli* (Chumpolkulwong et al. 2006). The translation efficiency of short ORF1 in pO-RFP was adjusted successfully using the inserted rare codons, and it was well reflected indirectly from the fluorescent signal derived from the downstream ORF2. Interestingly, due to the introduction of three consecutive rare codons into the upstream short ORF, the translation of mCherry in the downstream ORF was shut down in BL21(DE3)pLysS, but was partially recovered in Rosetta(DE3)pLysS (rare tRNAs were supplemented). Furthermore, the model protein eGFP was inserted into the upstream ORF1, and various expression levels of eGFP were obtained by varying rare codons preceding *egfp*. Importantly, a positive correlation was found between the expression levels of eGFP in ORF1 and mCherry in ORF2.

Fluorescence proteins have been successfully overexpressed in a wide range of organisms without obvious cytotoxic effects, and they have great potential as reporters for monitoring gene expression (Cubitt et al. 1995). In this study, no inclusion bodies were formed during over expression of mCherry in BL21(DE3)pLysS/pET-RFP (data not shown). Growth retardation and bacterial cell disruption were sometimes observed when recombinant proteins were expressed at high levels (Mahalik et al. 2014). Interestingly, red fluorescent signal was also detected in culture supernatant of BL21(DE3)pLysS/pET-RFP at the late phase of IPTG induction (data not shown). This suggests that the integrity of microbial cell factories could be checked during fermentation by determination of the fluorescence in culture supernatant.

Based on the strategy mentioned in this study, the fluorescent signal can be noninvasively detected in living bacterial cells during fermentation in real time, and it eliminates the analysis of exogenously expressed protein by routine procedures. There are still some shortcomings in the dicistronic system in this study. First of all, although no mCherry expression lag was observed in BL21(DE3)pLysS/pET-RFP, slight mCherry expression lag (about 1 h) still existed in the dicistronic vectors in this study. Compared to an enhanced GFP variant (eGFP), mCherry matures quickly at 37 °C with maturation half time of 15 min (Shaner et al. 2004). Previous studies showed both mCherry and eGFP could rapidly fold and mature even when fused to poorly folding polypeptides (Aper et al. 2014; Tsiens 1998). There was not enough evidence showing that 1 h lag of fluorescent signal resulted from incomplete maturation of mCherry. The expression of two ORFs involved in the dicistronic vectors in the study might be initiated in order, so there was a time lag between translation initiation of ORF1 and ORF2. In addition, most ribosomes dissociated at the stop codon of ORF1, followed by a great decrease in the expression of mCherry in ORF2. Due to the reasons above, the accumulation of matured mCherry might be below the detection limit at the early stage of induction. However, it hardly exerted any influence on the result after 2 h induction. More importantly, a 4–6 h induction was generally adopted in most protein expression studies. Second of all, it is inevitable that the synthesis of reporter mCherry in downstream ORF2 may lead to an unnecessary metabolic burden on the cell. Extra substrate and energy consumption might exert a negative influence on the yield of highly expressed target protein in ORF1, but the influence could also become insignificant when a poorly expressed target gene is inserted in ORF1. To decrease the uncertainties of the influence on the production of the target protein, a low molecular weight and easy detectable reporter might be a good solution.

#### Abbreviations

*E. coli*: *Escherichia coli*; LB: Luria–Bertani; RFP: red fluorescent protein; GFP: green fluorescent protein; IPTG: isopropyl- $\beta$ -D-thiogalactopyranoside; RBS: ribosome-binding site; MCS: multiple cloning sites; ORF: open reading frame; SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis; ELISA: enzyme linked immunosorbent assay; kDa: kilodalton; SD: Shine–Dalgarno.

#### Authors' contributions

This study was carried out in collaboration between all authors. CYH and XQH conceived and designed the experiments. YG and WZ performed the study. CYH drafted the manuscript. All authors read and approved the final manuscript.

#### Acknowledgements

We thank Anna Liu from Yale University for English language editing with an earlier version of the manuscript.

**Competing interests**

The authors declare that they have no competing interests.

**Availability of data and materials**

All data generated or analyzed during this study are included in this published article.

**Ethics approval and consent to participate**

Not applicable.

**Funding**

This work was supported by Natural Science Foundation of Guangdong Province (2015A030313838), Science and Technology Program of Shenzhen (JCYJ20150403091305481, JCYJ20160429090813380).

**Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 9 March 2018 Accepted: 12 May 2018

Published online: 21 May 2018

**References**

- Aper SJ, van Spreuwel AC, van Turnhout MC, van der Linden AJ, Pieters PA, van der Zon NL, de la Rambelje SL, Bouten CV, Merckx M (2014) Colorful protein-based fluorescent probes for collagen imaging. *PLoS ONE* 9(12):e114983
- Blumenthal T (2004) Operons in eukaryotes. *Brief Funct Genomic Proteomic* 3(3):199–211
- Brule CE, Grayhack EJ (2017) Synonymous codons: choose wisely for expression. *Trends Genet* 33(4):283–297
- Carrier TA, Keasling JD (1999) Library of synthetic 5' secondary structures to manipulate mRNA stability in *Escherichia coli*. *Biotechnol Prog* 15(1):58–64
- Cheng Y, Liu Z, Zeng J, Cheng L, Yan Z, Duan S, Feng X, Zheng K, Zheng X, Wang R (2016) Construction and co-expression of polycistronic plasmids encoding bio-degumming-related enzymes to improve the degumming process of ramie fibres. *Biotechnol Lett* 38(12):2089–2096
- Chumpolkulwong N, Sakamoto K, Hayashi A, Iraha F, Shinya N, Matsuda N, Kiga D, Urushibata A, Shirouzu M, Oki K, Kigawa T, Yokoyama S (2006) Translation of 'rare' codons in a cell-free protein synthesis system from *Escherichia coli*. *J Struct Funct Genomics* 7(1):31–36
- Cubitt AB, Heim R, Adams SR, Boyd AE, Gross LA, Tsien RY (1995) Understanding, improving and using green fluorescent proteins. *Trends Biochem Sci* 20(11):448–455
- de Smit MH, van Duin J (2003) Translational standby sites: how ribosomes may deal with the rapid folding kinetics of mRNA. *J Mol Biol* 331(4):737–743
- Espipov RS, Makarov DA, Stepanenko VN, Miroshnikov AI (2016) Development of the intein-mediated method for production of recombinant thymosin beta4 from the acetylated in vivo fusion protein. *J Biotechnol* 228:73–81
- Gasch AP, Eisen MB (2002) Exploring the conditional coregulation of yeast gene expression through fuzzy k-means clustering. *Genome Biol* 3(11):RESEARCH0059
- Gaymard E, Franchini L, Manieri W, Stutz E, Schurmann P (2000) A dicistronic construct for the expression of functional spinach chloroplast ferredoxin:thioredoxin reductase in *Escherichia coli*. *Plant Sci* 158(1–2):107–113
- Hui C, Guo Y, Zhang W, Gao C, Yang X, Chen Y, Li L, Huang X (2018a) Surface display of PbrR on *Escherichia coli* and evaluation of the bioavailability of lead associated with engineered cells in mice. *Sci Rep* 8(1):5685
- Hui CY, Guo Y, Yang XQ, Zhang W, Huang XQ (2018b) Surface display of metal binding domain derived from PbrR on *Escherichia coli* specifically increases lead(II) adsorption. *Biotechnol Lett* 40(5):837–845
- Isaacs FJ, Dwyer DJ, Ding C, Pervouchine DD, Cantor CR, Collins JJ (2004) Engineered riboregulators enable post-transcriptional control of gene expression. *Nat Biotechnol* 22(7):841–847
- Ito W, Kurosawa Y (1992) Development of a prokaryotic expression vector that exploits dicistronic gene organization. *Gene* 118(1):87–91
- Jacob F, Perrin D, Sanchez C, Monod J, Edelman S (2005) The operon: a group of genes with expression coordinated by an operator. *C.R. Acad. Sci. Paris* 250 (1960) 1727–1729. *C R Biol* 328(6):514–520
- Kane JF (1995) Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Curr Opin Biotechnol* 6(5):494–500
- Kudla G, Murray AW, Tollervey D, Plotkin JB (2009) Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* 324(5924):255–258
- Levin-Karp A, Barenholz U, Bareia T, Dayagi M, Zelcbuch L, Antonovsky N, Noor E, Milo R (2013) Quantifying translational coupling in *E. coli* synthetic operons using RBS modulation and fluorescent reporters. *ACS Synth Biol* 2(6):327–336
- Madje K, Schmolzer K, Nidetzky B, Kratzer R (2012) Host cell and expression engineering for development of an *E. coli* ketoreductase catalyst: enhancement of formate dehydrogenase activity for regeneration of NADH. *Microb Cell Fact* 11:7
- Mahalik S, Sharma AK, Mukherjee KJ (2014) Genome engineering for improved recombinant protein expression in *Escherichia coli*. *Microb Cell Fact* 13:177
- Nguyen VD, Hatahet F, Salo KE, Enlund E, Zhang C, Ruddock LW (2011) Pre-expression of a sulfhydryl oxidase significantly increases the yields of eukaryotic disulfide bond containing proteins expressed in the cytoplasm of *E. coli*. *Microb Cell Fact* 10:1
- Palleja A, Garcia-Valve S, Romeu A (2009) Adaptation of the short intergenic spacers between co-directional genes to the Shine-Dalgarno motif among prokaryote genomes. *BMC Genomics* 10:537
- Parmley JL, Huynh MA (2009) Clustering of codons with rare cognate tRNAs in human genes suggests an extra level of expression regulation. *PLoS Genet* 5(7):e1000548
- Peranen J, Rikkonen M, Hyvonen M, Kaariainen L (1996) T7 vectors with modified T7lac promoter for expression of proteins in *Escherichia coli*. *Anal Biochem* 236(2):371–373
- Pfleger BF, Pitera DJ, Smolke CD, Keasling JD (2006) Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nat Biotechnol* 24(8):1027–1032
- Rosenberg AH, Goldman E, Dunn JJ, Studier FW, Zubay G (1993) Effects of consecutive AGG codons on translation in *Escherichia coli*, demonstrated with a versatile codon test system. *J Bacteriol* 175(3):716–722
- Salis HM, Mirsky EA, Voigt CA (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* 27(10):946–950
- Shaner NC, Campbell RE, Steinbach PA, Giepmans BN, Palmer AE, Tsien RY (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat Biotechnol* 22(12):1567–1572
- Shetty RP, Endy D, Knight TF Jr (2008) Engineering BioBrick vectors from BioBrick parts. *J Biol Eng* 2:5
- Tsien RY (1998) The green fluorescent protein. *Annu Rev Biochem* 67:509–544
- Wang B, Wang Z, Yan J, Wang L, Wang Z, Wu J, Zhang H, Wu H, Kong W, Yu B, Yu X (2015) Expression, purification and characterization of heterotrimeric forms of sTRAIL using a polycistronic expression vector. *Protein Expr Purif* 115:118–124
- Wang Y, Li C, Khan MR, Wang Y, Ruan Y, Zhao B, Zhang B, Ma X, Zhang K, Zhao X, Ye G, Guo X, Feng G, He L, Ma G (2016) An engineered rare codon device for optimization of metabolic pathways. *Sci Rep* 6:20608
- Yadav K, Kumar C, Archana G, Kumar GN (2014) Artificial citrate operon and *Vitreoscilla* hemoglobin gene enhanced mineral phosphate solubilizing ability of *Enterobacter hormaechei* DHRSS. *Appl Microbiol Biotechnol* 98(19):8327–8336