Full Length Research Paper

Development of stable marker-free nuclear transformation strategy in the green microalga Chlorella vulgaris

Ahmed E. Gomma1*, Sung-Kwon Lee2*, Sang Mi Sun1, Seung Hwan Yang2,3* and Gyuhwa Chung1*

1Department of Biotechnology, Chonnam National University, Chonnam 550-749, Republic of Korea.
2Center for Nutraceutical and Pharmaceutical Materials, Myongji University, Gyeonggi 449-728, Republic of Korea.
3Interdisciplinary Program of Biomodulation, Myongji University, Yongin, Gyeonggi-Do, 449-728, Republic of Korea.

Received 22 April, 2015; Accepted 14 September, 2015

Although microalgae have valuable resources with great necessity in many of biotechnological applications, no tools have been developed yet for a stable genetic transformation without antibiotic marker genes in these organisms. Chlorella is one of the most useful genus for biotechnological applications. The transfer of foreign DNA (vector or linear DNA cassette) into Chlorella by electroporation has very low stability and it is hard to screen the transformants without antibiotic marker genes. However, the marker genes have some disadvantages to host cells. To avoid the negative effects caused by the marker genes, we tried to develop a stable marker-free nuclear transformation system in Chlorella. For this, linear gene expression cassettes (LGEC) were constructed with functional domains, which are responsible for transformation, of SV40 large T antigen. The LGECs were transferred into Chlorella via electroporation and durability of the LGECs were confirmed in transgenic Chlorella. Transcription levels of the LGECs were also determined at different cell cycle sates. The LGECs integrated into the chromosomal DNA of Chlorella were stably replicated and were expressed successfully at G0-, G1-, and G2-phases. This study presents a stable marker-free nuclear transformation system with potential for biotechnological applications.

Key words: Chlorella vulgaris, marker-free nuclear transformation, SV40 large T antigen, microalga.

INTRODUCTION

In recent years, microalgae have become valuable resources with its great necessity in many of biotechnological applications (Chu, 2012). Currently, to improve microalgae for industrial application, various
Figure 1. Schematic diagram of functional domains of SV40 large T antigen.

genesic engineering approaches have been tried (Hlavová et al., 2015). Generally, generating transgenic microalgae via genetic transformation requires antibiotic marker genes to increase stability of a foreign gene expression cassette (Niu et al., 2011). However, most antibiotic marker genes have some disadvantages on the host cell. Many algae have natural resistance to most antibiotics or herbicides commonly used for genetic transformation and the stress caused by antibiotics and expression of antibiotic marker genes might negatively affect growth and metabolism of host cell (Ahuja et al., 1996; Niu et al., 2011). So, it is very necessary to develop a stable marker-free nuclear transformation system in microalgae.

Simian virus 40 (SV40) is a member of Polyomaviridae and its genome encodes three proteins that are structural components of the virion (VP1, VP2, VP3), and two nonstructural proteins namely large T antigen (T antigen) and small T antigen (t antigen) (Ahuja et al., 2005). The SV40 large T antigen is a viral initiator protein for DNA replication. It is a sequence specific DNA binding protein that also possesses intrinsic DNA helicase and ATPase activities, besides nuclear localization signal (Figure 1). These activities of large T antigen are required for its ability to influence viral DNA replication (Ahuja et al., 2005). The previous studies reported the functions of SV40 T antigen on the transformation of mammalian and yeast cell (Ali and DeCaprio, 2001; Sáenz-Robles et al., 2001; Sullivan and Pipas, 2002). The SV40 large T antigen contains three different domains, which contribute to cellular transformation. These domains include the J domain for interaction with the heat shock chaperone, Hsc70; the LXCXE motif responsible for binding the retinoblastoma (Rb) family of tumor repressors, such as p107 and p130; the bipartite domain to interact with the cellular tumor suppressor protein, p53 (Sullivan et al., 2000; 2001; Sáenz-Robles et al., 2001; Ahuja et al., 2005). The interaction of SV40 small T antigen with the cellular phosphatase pp2A is also required for the transforming function of SV40 T antigen (Pallas et al., 1990; Hahn et al., 2002).

The Rb-E2F pathway is a major mechanism connecting the activity of the cell-cycle machinery to transcription (Verma and Hong, 2001). Recently, Rb-E2F cell-cycle signaling pathway has been recognized in plants like in animals, and it is conserved in plants (De Veylder et al., 2007; Gutierrez, 2009), and several studies revealed that microalgae also have the retinoblastoma (Rb) protein homologue (Cross and Roberts, 2001; Umen and Goodenough, 2001; Kianianmomeni et al., 2008; Moulager et al., 2010). Moreover, the previous reports demonstrated that the SV40 genomic DNA was able to easily integrate into the host chromosomal DNA via nonhomologous recombination with nuclear localization signal (NLS) (Campos et al., 1978; Hwang and Kucherlapati, 1983; Dorsett et al., 1985; Neuhaus et al., 1986; Ahuja et al., 2005). Based on these reports, we supposed that the SV40 T antigen might be applicable to develop the stable marker-free nuclear transformation system without antibiotic marker genes in microalgae.

In the present study, we constructed several linear gene expression cassettes (LGECs) for nuclear transformation in *Chlorella vulgaris* with specific parts of SV40 T antigen, such as NLS and DNA binding domain (DNA-BD), and then identified the durability of transformed LGECs in the genome of *C. vulgaris*.

**MATERIALS AND METHODS**

**Microalgae materials and culture conditions**

Freshwater sample was collected from a small pool next to Deokgok-je reservoir, Josan-ri, Soramyeon, Yeosu, South Korea (latitude 34° 46’ 47.9˝ N, longitude 127° 37’ 29.0˝ E). The algal strain was isolated from the fresh water sample, and was cultivated in Beijerinck medium. The isolated algal strain was identified as *C. vulgaris* by morphological identification and 18S rDNA sequencing. The purified *C. vulgaris* was grown in 30 ml *Euglena gracilis*; Jaworski 1:1 medium (EG:JM medium) for two weeks at 24°C and 80 rpm under continuous illumination (light intensity: 8,000 Lm).

**Construction of linear gene expression cassette**

The PCR primers used for linear gene expression cassette construction are listed in Table 1. SV40 genomic DNA (ATCC
Table 1. Primers used in this study.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR</td>
<td></td>
</tr>
<tr>
<td>35Sf</td>
<td>TCCCCAGATTAGCCTTTTCA</td>
</tr>
<tr>
<td>35Sr</td>
<td>TCTTTTTTGGCCCGTGTTCT</td>
</tr>
<tr>
<td>NLSf</td>
<td>AGAACACGGGGCAAAAAAGA</td>
</tr>
<tr>
<td>NLSr</td>
<td>CACAAGTGGATACCTTTTCT</td>
</tr>
<tr>
<td>BDf</td>
<td>GAGAAAGGTATCCACTTGTG</td>
</tr>
<tr>
<td>BDr</td>
<td>TAAAGATAAGTTAGCAATTCT</td>
</tr>
<tr>
<td>HSPf</td>
<td>GAATTGCTAACTTATCTTTA</td>
</tr>
<tr>
<td>HSPr</td>
<td>CCTATGGAATATGAAGATG</td>
</tr>
<tr>
<td>Orif</td>
<td>CATCCTCATATTCGGCCCC</td>
</tr>
<tr>
<td>Orir</td>
<td>GGCCTCTGCAATAAATAAAAA</td>
</tr>
<tr>
<td>18Sf</td>
<td>CTAGAGGAAGAGAAGTCGT</td>
</tr>
<tr>
<td>18Sr</td>
<td>ACGACTCTCTCTCTCTCTAG</td>
</tr>
<tr>
<td>35S2r</td>
<td>TCATTGGAGAGACACGCG</td>
</tr>
<tr>
<td>RT-PCR</td>
<td></td>
</tr>
<tr>
<td>BD2f</td>
<td>TTTAGAATGTGGTTTGGACT</td>
</tr>
<tr>
<td>BD2r</td>
<td>AGATGGCATTCTCTTGAGC</td>
</tr>
</tbody>
</table>

Figure 2. Structures of linear gene expression cassettes.

accession: 45019D) was used as a template source for the NLS, DNA binding domain, and replication origin (ori). CaMV 35S promoter and HSP-terminator were isolated from pRI 201 vector (Takara, Dalian, China). To construct LGECs, CaMV 35S promoter, NLS, DNA-BD, HSP-terminator, and 18S rRNA gene fragments were amplified by PCR with specific primer pairs: 35Sf/35Sr, NLSf/NLSr, BDf/BDr, HSPf/HSPr, and Orif/Orir. The structures of each construct are shown in Figure 2. All fragments for the LGECs were connected by the overlap extension PCR, and the detailed procedure is shown in Figure 3. For phosphorylation of the LGEC-A and -C, the primers 35Sf, 35Sr, 18Sf, and 18Sr were phosphorylated at 5’ end with T4 polynucleotide kinase (NEB, Hitchin, UK) and PCR amplifications were carried out with these primer pairs using LGEC-A or -B as a template. The overlapping PCR products were examined by electrophoresis on 1.5% agarose gels, and DNA bands of the correct size were extracted using the
Figure 3. Schematic flow diagram for the construction of naked marker-free DNA cassettes.

Figure 3. Schematic flow diagram for the construction of naked marker-free DNA cassettes.

QIAquick Gel Extraction Kit (Qiagen, CA, USA). All PCR reactions were carried out using the high fidelity DNA polymerase (TLA polymerase) (Bioneer, Daejeon, Republic of Korea).

**Generation of transgenic *C. vulgaris* by electroporation**

The LGECs described above were used to transform *C. vulgaris* by electroporation. *Chlorella* cells in the exponential phase were collected by centrifuge at 1,300 × g for 10 min. *Chlorella* cells were washed out with 20 ml of HEPES-glucose buffer (7 mM HEPES and 252 mM glucose, pH 7.0) two times. The equivalent of 1 × 10⁸ *Chlorella* cells was resuspended in 2 ml of sterilized distilled water and was kept on ice for 2 h. Suspension aliquots of 80 μl were mixed with 33.6 ng of each LGEC then transferred into an electroporation cuvette. Electroporation was carried out with a
GenePulser Xcell apparatus (Bio-Rad, CA, USA) according to the following conditions: electric field 1,800 V cm$^{-1}$ and shunt 200 Ω. After electroporation, cells were kept on ice for 5 min and resuspended in 5 ml EG:JM medium containing 1 M mannitol and 1 M sorbitol, and then incubated in the dark for 24 h. After incubation, 1 ml of suspension was transferred into 30 ml EG:JM medium and cultivated for 24 h under continuous illumination. Afterwards, cells were collected by centrifugation at 1300 × g for 5 min and resuspended in 1 ml EG:JM medium. Finally, the cells were spread on Petri dishes in EG:JM agar medium.

**Isolation of transgenic C. vulgaris**

After five days of incubation of the plates under illumination (16:8 LD), two colonies were randomly picked from each transgenic line and each colony was suspended in 20 μl distilled water in PCR tube, and then the tubes were placed in a boiling water for 10 min. Afterwards, PCR was carried out with CaMV 35S promoter specific primer pair (35Sf and 35Sr) using the boiled lysates as a template. The correct transgenic C. vulgaris was regenerated in 30 ml EG:JM medium for 14 days under continuous illumination with shaking. To calculate the transformation frequency for each LGEC, each LGEC was transferred into C. vulgaris by electroporation according to the method mentioned above. Then, 100 μl of sample was taken from the 1 ml of cell suspension and was spread on Petri dishes in EG:JM agar medium. After incubation, the transformation frequency was calculated with the following formula:

$$\text{Transformation frequency (\%)} = \frac{\text{Number of PCR positive colonies}}{\text{Total number of colonies}} \times 100$$

**Analysis of the durability and the transcriptional level of the naked DNA cassette in transgenic C. vulgaris**

To determine the durability of the LGECs in transgenic C. vulgaris, the transformants were serially subcultured three times with six days intervals at 24°C and 80 rpm in 30 ml EG:JM medium under continuous illumination, and the presence of LGEC was determined at each subculture cycle by PCR. Genomic DNA was extracted from transgenic C. vulgaris according to previously described procedures (Dawson et al., 1997). PCR analysis was performed using Ex-taq (Takara, Dalian, China) with specific primer pair (35Sf and 35Sr), which is designed to amplify 843-bp of CaMV 35S promoter in the LGECs. The transcription levels of LGECs in transgenic C. vulgaris was also analyzed by reverse-transcriptase PCR(RT-PCR). For RT-PCR analysis, transgenic C. vulgaris were grown in 30 ml of EG:JM medium and Chlorella cells were collected from 4 ml culture at each growth stage, G0-, G1-, and G2-phase. Total RNA was extracted using Isol-RNA Lysis reagent (Takara, Shiga, Japan). The first strand cDNA synthesized using Maxima First Strand cDNA synthesis Kit for RT-PCR (Thermo Scientific Fermentas, MD, USA) with 1 μg of total RNA as a template. PCR was performed with DNA binding domain specific primer pair: BD2f/BD2r, which was designed to amplify 21-bp of DNA fragment. To assess the genomic DNA contamination of the total RNA sample, PCR was carried out with the specific primers pair.

**Results**

**Construction of LGEC**

The 843-bp CaMV 35S promoter, 21-bp NLS, 384-bp DNA binding, 78-bp replication origin, 249-bp Hsp-terminator regions were amplified by PCR, and confirmed by sequencing analysis. All components were connected by overlap extension PCR to produce the LGEC-A. To add the 18S rDNA flank fragments at both ends of the cassette, PCR was carried out with the primer pair 18Sf/18Sr using the LGEC-A as a template. The resultant LGEC-A and -C were 1,575-bp and 1,615-bp, respectively. For the phosphorylation of the LGEC-A and -C, the primer pairs 35Sf/35Sr and 18Sf/18Sr were phosphorylated at 5'end, and then PCR was performed with these primer pairs using the LGEC-A or -C as a template. A schematic flow diagram of LGEC construction is shown in Figure 3.

**Selection of transgenic C. vulgaris**

Each LGEC was introduced into cells of C. vulgaris by electroporation. C. vulgaris transformants were regenerated in EG:JM medium for 24 h, and then cells were spread on agar medium. After 5 days, 2 colonies were randomly picked from each transgenic line and the presence of LGEC was determined by colony PCR with CaMV 35S specific primer pair. Figure 4 shows the result of PCR analysis. The DNA bands with exact size (843 bp) were amplified from eight colonies of transgenic C. vulgaris cells with LGEC-A, -B, -C, or -D and these results demonstrated that the LGEC were successfully transformed into C. vulgaris cells. The transformation frequency of each LGEC was also determined and shown in Table 2.

**Analysis of durability of LGEC in transgenic C. vulgaris**

To analyze durability of LGECs in transgenic C. vulgaris, each transgenic line was subcultured 3 times and presence of CaMV 35S promoter was verified by PCR (Figure 5). Samples of each subculture cycle were subjected to PCR analysis with CaMV 35S promoter specific primer pair, and the results demonstrate that 843-bp DNA fragments were amplified from all subculture cycle samples of transformed C. vulgaris cells.

**Expression analysis of transgenic C. vulgaris**

To verify the expression of each LGEC in transgenic C. vulgaris, transcription level of the gene encoding DNA binding domain of SV40 large T antigen was analyzed by RT-PCR. For the analysis of gene transcription, each
Figure 4. Colony PCR results of transformed *C. vulgaris*. Lanes: M, 100-bp DNA ladder; 1, PCR of colony number 1 of transgenic *C. vulgaris* with cassette A; 2, PCR of colony number 2 of transgenic *C. vulgaris* with cassette A; 3, PCR of colony number 1 of transgenic *C. vulgaris* with cassette B; 4, PCR of colony number 2 of transgenic *C. vulgaris* with cassette B; 5, PCR of colony number 1 of transgenic *C. vulgaris* with cassette C; 6, PCR of colony number 2 of transgenic *C. vulgaris* with cassette C; 7, PCR of colony number 1 of transgenic *C. vulgaris* with cassette D; 8, PCR of colony number 2 of transgenic *C. vulgaris* with cassette D; 843-bp bands in lane 1-8 indicate the expected CaMV 35S promoter PCR product.

Table 2. Transformation frequencies for the linear gene expression cassettes.

<table>
<thead>
<tr>
<th>Cassette</th>
<th>Frequency (% mean ± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>63.2 ± 3.1</td>
</tr>
<tr>
<td>B</td>
<td>62.3 ± 4.8</td>
</tr>
<tr>
<td>C</td>
<td>62.7 ± 2.4</td>
</tr>
<tr>
<td>D</td>
<td>61.6 ± 4.9</td>
</tr>
</tbody>
</table>

*Values are means of three replicate determination; SD, standard deviation.

Figure 5. PCR analysis of transgenic *C. vulgaris* in three generations. Each *C. vulgaris* line was subcultured 3 times with 6 days intervals and the presence of CaMV 35S promoter was determined by PCR. The DNA bands indicate the PCR product of 843-bp sized CaMV 35S promoter.
transgenic line of *C. vulgaris* was cultivated in EG:JM medium, and aliquots of algal culture were collected at G0-, G1-, and G2-phases. Total RNA was isolated from each sample and subjected to RT-PCR. Figure 6 shows that each LGEC was successfully expressed in transgenic *C. vulgaris* at G0-, G1- and G2-phases. The density of PCR band was calculated using Image J software (http://rsb.info.nih.gov/ij).

### Analysis of cell growth

The wild-type and the transgenic lines of *C. vulgaris* were cultivated in EG:JM medium for 8 days and the optical density was measured at 680 nm. No significant differences in cell growth were observed among the wild-type and the transgenic *C. vulgaris* lines (Figure 7).

### DISCUSSION

In genetic engineering of microalgae, antibiotic marker genes have been widely used for the selection of transformed cells and for the maintenance of foreign vector or DNA cassette. However, antibiotic marker genes might cause negative effects on host cells, such as a decline in growth rate and a decrease in production of secondary metabolite (Apt et al., 1996; Niu et al., 2011). To avoid disadvantages of antibiotic marker genes, several methods, such as co-transformation, transposition, and site-specific recombination, have been developed to exclude marker genes from transgenic plants (Nishizawa-Yokoi et al., 2014). Moreover, marker-free chloroplast transformation system has successfully been developed in the green microalga *Chlamydomonas reinhardtii* (Chen and Melis, 2013). However, to our knowledge, there were no report about the nuclear transformation method without using antibiotic marker genes in microalgae *C. vulgaris*.

In order to develop the stable marker-free nuclear transformation system in *C. vulgaris*, we tried to enhance the durability of gene expression cassette via increasing the concentration of gene expression cassette in nulear.
Figure 7. Cell growth analysis of the wild-type and the transgenic *C. vulgaris*. **A.** The wild-type *C. vulgaris*. **B.** Transgenic *C. vulgaris* with cassette A. **C.** Transgenic *C. vulgaris* with cassette B. **D.** Transgenic *C. vulgaris* with cassette C. **E.** Transgenic *C. vulgaris* with cassette D. Results represent the average of three replicates (*n* = 3). Error bars represent standard deviations.

To achieve this, the LGECs including the responsible part for cell transformation in the SV40 large T antigen were constructed (Zalvide et al., 2001; Ahuja et al., 2005; White and Khalili, 2006). Additionally, to enhance the durability and the integration rate of LGEC, the modified LGECs were produced by 5' end phosphorylation or addition of 18S rDNA flank fragments. Each LGEC was transferred into the *C. vulgaris*, and integration of the cassette in transgenic *C. vulgaris* was confirmed by colony PCR. Then, the PCR analysis displayed that all LGECs were successfully integrated into the genomic DNA of *C. vulgaris*. Additionally, the transformation frequency of each LGEC was also calculated, and there were no significant differences in the transformation frequency observed among the LGECs.

The durability of LGECs in transgenic *C. vulgaris* was also confirmed in three generations, and the PCR analysis results showed that the LGECs were stably maintained after three times of subculture. Additionally, no significant differences in durability were observed among four LGECs, and these results demonstrated that the addition of 18S rDNA flank fragments and phosphate groups at both ends of the LGEC did not affect the durability of the cassettes.

The functionality of the SV40 large T antigen based marker-free expression system was also confirmed via analysis of LGEC expression levels. For transcription analysis, transgenic *C. vulgaris* cells were collected at G0-, G1- and G2-phases, and transcription level of DNA binding domain was analyzed by RT-PCR. Transcription analysis demonstrated that all LGECs were expressed at three cell cycle states. Expression level of LGECs was highest at G0-phase, and then they were decreased gradually from the G1 to the G2-phase. There were no
differences in the expression levels among four LGECs at G0- and G2-phases, however, significant differences were observed at G1-phase. The LGEC-D showed highest expression level at G1-phase and 2.8-, 1.7- and 3.3-fold was higher compared to LGEC-A, -B, and -C, respectively. However, we could not predict the reason for this phenomenon and it was supposed that the combination of 18S rDNA flank fragments and 5’ phosphorylation might affect the expression of foreign DNA cassette. Additionally, to determine the effect of LGECs on cell growth, the cell growth of the wild-type and the transgenic C. vulgaris was observed, and no significant differences in cell growth were observed among the wild-type and the transgenic C. vulgaris lines. This result demonstrate that the linear gene expression cassette using SV40 large T antigen components have no negative effect on the growth of C. vulgaris.

Conclusion

This study shows that the functional domains for host cell infection derived from SV40 large T antigen are working in the microalga Chlorella. The linear gene expression cassettes include SV40 large T antigen components which can be stably expressed in the Chlorella cells without using antibiotic marker genes. To our knowledge, this is the first report of the marker-free nuclear transformation in microalgae and may contribute to the development of genetically engineered microalgae. Further studies are needed to rapidly isolate the transgenic microalgae.

Conflict of interests

The authors did not declare any conflict of interest.

ACKNOWLEDGEMENT

This work was carried out with the support of "Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ01131501)" Rural Development Administration, Republic of Korea.

REFERENCES