Robust expression of heterologous genes by selection marker fusion system in improved *Chlamydomonas* strains

Fantao Kong,[1] Tomohito Yamasaki,[1] Sari Dewi Kurniasih,[1] Liyuan Hou, [1] Xiaobo Li,[2] Nina Ivanova,[2] Shigeru Okada,[3,4] and Takeshi Ohama[1*]

[1]School of Environmental Science and Engineering, Kochi University of Technology (KUT), Tosayamada, Kochi 782-8502; [2]Carnegie Institution for Science 260 Panama St. Stanford, CA 94305, USA; [3]Laboratory of Aquatic Natural Products Chemistry, Graduate School of Agricultural & Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan; and [4]Japan Science and Technology Agency (JST), 5 Sanbancho, Chiyoda, Tokyo 102-0075, Japan

Keywords:

transgene silencing, position effect, squalene synthase-like gene, *Botryococcus braunii*, methyltransferase gene

Corresponding author:

Takeshi Ohama

E-mail: ohama.takeshi@kochi-tech.ac.jp.

Tel.: +81-887-57-2512

Fax: +81-887-57-2520

Short title: Heterologous expression of squalene synthase-like gene

Abstract

Chlamydomonas is a very attractive candidate plant cell factory. However, its main drawback is the difficulty to find the transformants that robustly express heterologous genes randomly inserted in the nuclear genome. We previously showed that domestic squalene synthase (*SQS*) gene of *Chlamydomonas* was much more efficiently overexpressed in a mutant strain [UV-mediated mutant (UVM) 4] than in wild type. In this study, we evaluated the possibility of a new mutant strain, *met1*, which contains a tag in the maintenance type methyltransferase gene that is expected to play a key role in the maintenance of transcriptional gene silencing. The versatile usefulness of the UVM4 strain to express heterologous genes was also analyzed. We failed to overexpress *CrSSL3* cDNA, which is the codon-adjusted squalene synthase-like gene originated from *Botryococcus braunii*, using the common expression cassette in the wild-type CC-1690 and UVM4 strains. However, we succeeded in isolating western blot-positive transformants through the combinational use of the UVM4 strain and ble2A expression system of which expression cassette bears a fused ORF of the target gene and the antibiotic resistance gene *ble* via the foot-and-mouth disease virus (FMDV) self-cleaving 2A sequence. It is noteworthy that even with this system, huge deviations in the accumulated protein levels were still observed among the UVM4 transformants.

Introduction

Recently, microalgae have attracted much attention as plant cell factories for the production of various commercial products, including biofuels, pharmaceutical terpenoid, nutraceuticals, and therapeutics (1-3). However, until now, only limited transgenic products have been commercialized mainly because of the difficulty in improving strains for abundant accumulation of a product of interest above the required levels.

The unicellular green alga, *Chlamydomonas reinhardtii*, has been a prominent model organism for genetic studies primarily because of the efficient genetic transformation techniques for mitochondria, chloroplasts, and nucleus (4-7). The genome (nuclear, plastid, and mitochondria) of *C. reinhardtii* has been fully sequenced (8), and large chemical and insertional mutant libraries have been established. On the other hand, the major disadvantage of *C. reinhardtii* is the poor expression of transgenes from the nuclear genome (9, 10). The molecular mechanism(s) of this is still uncovered, and a possible reason for this is closely related to the strong transcriptional silencing against transgenes (2, 11), which is mediated by both DNA methylation and DNA methylation-independent pathways (12, 13).

Over the past two decades, several advanced strategies have been developed to improve the expression of transgenes in *Chlamydomonas*, i.e., codon optimization (9, 10), utilization of endogenous intron(s) (14), and development of artificial strong promoter (15). Recently, UV-mediated mutant (UVM) strains of *Chlamydomonas* have been isolated for improved transgene expression. The UVM strains harbor unknown and unmapped mutation(s), and it has been suggested that their epigenetic transgene suppression mechanisms have been successfully knocked out (16).

Moreover, a novel nuclear expression system was reported to robustly express

heterologous genes (2). The system utilizes the foot-and-mouth disease virus (FMDV) 2A "self-cleaving" peptide to transcriptionally fuse a transgene open reading frame (ORF) to the antibiotic resistance marker gene *ble* (referred to as ble2A system in this paper). The FMDV 2A peptide, which is a short peptide with approximately 20 amino acid sequences, mediates ribosome-skipping reaction during translation (17). Because of this reaction, when 2A is fused between two ORFs to generate a single ORF, the resulting products are the two discrete proteins with the short 2A peptide sequence fused to the C-terminus of the first protein product, whereas the following protein has only one amino acid of the peptide covalently attached to the N-terminus. The efficient ability of the ble2A system to improve the heterologous expression of transgenes compared with the traditional nuclear expression vector in which the expression of the interest and selection marker genes controlled under independent promoters has been reported (2, 18).

Previously, we demonstrated that the *Chlamydomonas* squalene synthase (*SQS*) cDNA was much more efficiently expressed in the UVM4 and UVM11 than in wild-type strains (19). Recently, using the *Chlamydomonas* MmeI-based insertion site Sequencing (ChlaMmeSeq) method (20), an insertional mutant of *Cre10.g461750* was isolated. This gene encodes DNA methyltransferase 1 (*Dnmt1*) (20), which is expected to be involved in the maintenance of DNA methylation patterns (21, 22). Robust transcriptional gene silencing through DNA methylation is one of the major pathways for stable repression of transgenes. In this study, to expand the platform for efficient expression of various transgenes, we evaluated the potency of this tag-inserted strain by comparing domestic *SQS* gene expression levels in four strains CC-124 (wild-type), UVM4, UVM11, and the insertional mutant of *Cre10.g461750* ("*met1*" see below). Moreover, to evaluate the ability of the ble2A system for

overexpression of codon-optimized transgenes, a codon-adjusted *SQS*-like 3 gene (*CrSSL3*) and *CrSSL1*, which originated in *Botryococcus braunii*, were heterogeneously expressed in wild-type strains and in the UVM4 strain. These *SSL* genes are key enzymes for the biosynthesis of botryococcene in *B. braunii* B-race (23). The expression levels of these enzymes were closely analyzed by western blotting using a monoclonal antibody against gp-64 epitope.

Materials and Methods

Construction of the transformation vectors

The construction of the transformation vector containing *SQS* expression cassette was shown in detail by Kong et al. (19). For the construction of the P_{AR4}::ble-2A-SSL::term expression vector, the *ble* sequence, which contained one copy of the *RbcS2* intron 1, was fused in frame to the codon-optimized FMDV 2A coding sequence (2), and synthesized as an *Xba*I-*Nde*I/*Kpn*I fragment. The ble-2A fragment was inserted into a pSTBlue-1 plasmid (EMD Biosciences, USA) as *Xba*I/*Kpn*I fragment, generating the recombinant plasmid pSTBlue-1-ble-2A. The *Hsp70A* promoter fragment was amplified by polymerase chain reaction (PCR) with high-fidelity PrimeSTAR HS DNA polymerase (Takara, Japan) from the pALM32 plasmid (24) using primers *Xba*I-Hsp70A-F (5′-AATCTAGAGACGGCGGGG-3′) and *Nde*I-*Hind*III-Hsp70A-R (5′-CATATGAACTGAAGCTTGAGTGGTTATGTA-3′). This fragment was inserted into the pSTBlue-1-ble-2A plasmid as a *Xba*I/*Nde*I fragment, generating the recombinant plasmid pHsp70A-ble-2A.The fragment containing the sequence of *RbcS2* 3′ untranslated region (UTR) terminator was excised from the pHsp70A/RbcS2-cgLuc plasmid (25) by *BamH*I-*Kpn*I digestion and cloned into pHsp70A-ble-2A, resulting in the recombinant plasmid pHsp70A-ble-2A- term. For

the construction of four parallel copies of the first intron (intron 1) of *RbcS2*, the sense and antisense single-stranded oligonucleotides intron 1-left (5′-CAGGTGAGTCGACGAGCAAGCCCGGCGGATCAGGCAGCGTGCTTGCAGATTTGACTTGCAACGCCCGCATTGTGTCGACGAAGGCTTTTGGCTCCTCTGT-3′) and intron 1-right (5′-TGCCTGCAGGAATTCGATTGGTCTTGGCCATCCTGCAAATGGAAACGGCGACGCAGGGTTAGATGCTGCTTGAGACAGCGACAGAGGAGCCAAAAGCCTT-3′), respectively, were synthesized, annealed, and used as the templates to generate the fragment of intron 1 using the following primers: intron 1-left-F (5′-AAGCTTGATTGTCATGGCCAGGTGAGTCGACGAGCAAG-3′) and intron 1-right-R (5′-CCATGGGATATCGCATGCCTGCAGGAATTCGATTG-3′) by employing the overlap extension PCR (OE-PCR) (26) method. The fragment was then used as the template to amplify four parallel copies of intron 1 using the following primers: *Hind*III-c1-F (5′-AAGTAAAAGCTTGATTGTCATGGCCAG-3′) and *Sac*I-c1-R (5′-AAGTAAGAGCTCCCATGGGATATCGCATGC-3′) for intron 1-copy 1; *Sac*I-c2-F (5-AAGTAAGAGCTCGATTGTCATGGCCAGGTG-3′) and *Xba*I-c2-R (5′-AAGTAATCTAGACCATGGGATATCGCATGC-3′) for intron 1-copy 2; *Xba*I-c3-F (5′-AAGTAATCTAGAGATTGTCATGGCCAGGTG-3′) and *Sac*I-c1-R for intron 1-copy 3; *Sac*I-c2-F and *Kpn*I-c4-R (5′-AAGTAAGGTACCCCATGGGATATCGCATGC-3′) for intron 1-copy 4. These four fragments of copies of intron 1 were double digested using the introduced restriction enzymes and then ligated using the Mighty Mix DNA ligation kit (Takara) to generate four parallel copies of intron 1 of *RbcS2* as *Hind*III/*Kpn*I fragment. The *RbcS2* promoter fragment was generated by PCR from pHsp70A/RbcS2-cgLuc plasmid using the primers *Kpn*I-RbcS2-Pro-F2 (5′-TAAGGTACCCCGGGCGCGCCA-3′) and

*Nde*I-RbcS2-Pro-R2 (5′-CTTGGCCATATGTTTAGATGTTGAGTGACT-3′). The obtained fragment containing four copies of intron 1 and *RbcS2* promoter were digested by the *Hind*III-*Kpn*I and *Kpn*I-*Nde*I restriction enzyme couples, respectively, and then inserted into the *Hind*III/*Nde*I sites of pHsp70A-ble-2A-RbcS2 plasmid, generating the recombinant plasmid PAR4::ble-2A::term expression vector.

The codon usage of *CrSSL3* and *CrSSL1* cDNAs was optimized for *Chlamydomonas* (http://www.kazusa.org.jp/codon), and the codon-adjusted 1 × gp64-tag (27) sequence was attached at the N-terminus. These sequences were synthesized as *Xho*I/*Kpn*I fragments. For the construction of the GS-linker-3 × gp64-tag fragment, the sense and antisense single-stranded oligonucleotides GS-linker-1 × gp64-tag (5′-ATGGGCGGCAGCGGCGGCGGCAGCGGCGGCGGCAGCGGCTCCTGGAAGG ACGCGAGCGGCTGGAGCATCAGCGGCTCCTGGAAGGA-3′) and 2 × gp64-tag (5′-CGTGCCCTCAGTGGATCCTTATTAGCTCCAGCCGCTCGCGTCCTTCCAGGA GCCGCTCCAGCCGCTCGCGTCCTTCCAGGAGCCGCTGAT-3′) were synthesized, annealed, and used as the templates to generate the *Kpn*I-GS-linker-3 × gp64-tag-*BamH*I fragment by employing the OE-PCR method with the following primers: *Kpn*I-GS-linker-3 × gp64-tag-F (5′-AAGTAAGGTACCATGGGCGGCAGCGGC-3′) and *BamH*I-3 × gp64-tag-R (5′-CGTGCCCTCAGTGGATCCTTATTA-3′). The fragments of the *SSL3* and *SSL1* cDNA cassettes and GS-linker-3 × gp64-tag were double digested by the *Xho*I-*Kpn*I and *Kpn*I-*BamH*I restriction enzyme couples, respectively, and then cloned into the PAR4::ble-2A::term vector as an *Xho*I/*BamH*I fragment to generate the *SSL3* transformation vector that is schematically shown in Fig. 1. The DNA sequences of all the constructs were confirmed by direct sequencing using the dideoxy chain

termination method (28).

C. reinhardtii strains, growth, and transformation conditions

The *C. reinhardtii* strains CC-124 (wild type, mt^-) and CC-1690 (wild type, mt^+) were provided by the *Chlamydomonas* Resource Center (Minnesota, USA), the C-9 strain (wild type, mt^-, NIES-2235) was provided from the National Institute for Environmental Studies (NIES, Japan), and the UVM4 and UVM11 strains (16) were kindly provided by Dr. R. Bock (MPI-MP, Germany). Unless otherwise stated, cells were cultivated mixotrophically at 25°C in Tris-acetate phosphate (TAP) medium (29) under moderate and constant white fluorescent light (84 μmol photons m^{-2} s^{-1}) with gentle shaking. Nuclear transformation was performed using the electroporation method (30). Briefly, the cells were grown to 1.0–2.0 × 10^6 cells/mL in TAP medium. Subsequently, 2.5 × 10^7 cells were harvested by centrifugation and suspended in 250 μL of TAP medium supplemented with 50 mM sucrose (TAP/sucrose). Electroporation was performed by applying an exponential electric pulse of 0.7 kV at a capacitance of 50 μF (BTX, USA), using 300 ng of linearized plasmids purified by agarose gel electrophoresis according to the manufacturer's instructions. The transgenic strains were selected directly on TAP/agar plates containing zeocin (15 mg/L), and the plates were incubated under continuous fluorescent light (20 μmol m^{-2} s^{-1}) at 25°C.

PCR screening and analyses of the transformants

Primary antibiotic-resistant transformants were screened for the presence of the expression cassettes by PCR as described previously (31, 32). In brief, cells (1.0–5.0 × 10^6) were resuspended in Tris–EDTA solution and incubated at 98°C for 10 min.

Aliquots (1 μL) of the supernatants from denatured cell lysates were then used as template for 20 μL PCR, using promoter-specific forward and gene-specific reverse primers for *SQS* as previously described (19), and gene-specific forward primer (5′-AGATGGAGGCCAAGTGCGTC-3′) and terminator-specific reverse primer (5′-CCGCTTCAGCACTTGAGAGCA-3′) for *SSL3*. The amplification conditions were as follows: 98°C for 5 min; followed by 30 cycles at 95°C for 15 s, 58°C for 30 s, and 72°C for 30 s; and a final step at 72°C for 7 min using a Thermal Cycler 2720 (Applied Biosystems).

For semi-quantitative reverse transcription PCR, total RNA was isolated as previously described (34) using TRIzol reagent (Molecular Research Center, http://www.mrcgene.com/), according to the manufacturer's instructions. The purified total RNA was treated with DNase I (Takara, Japan) to remove residual genomic DNA contamination. First-strand cDNA was synthesized using an oligo(dT)18 primer or random heptamers and PrimeScript reverse transcriptase (Takara) according to the manufacturer's instructions. The cDNA fragment of *CrSSL1* was amplified by PCR using a set of primers: *CrSSL1*-F (5′-ATGACTATCAAGCGCCTGCAGAG-3′), and *CrSSL1*-R (5′-CCGCTTCAGCACTTGAGAGCA-3′).

Western blot screening of the transformants

Total cell extracts and soluble fractions were prepared as previously described (19, 33). For immunoblot analysis, proteins were separated on 12% SDS-PAGE and immunoblotted with the monoclonal anti-baculovirus envelope gp64 polypeptide antibody (1/5,000) (eBioscience, CA, USA) for detecting the gp64-tagged proteins, and anti-Histone H3 antibody (1/10,000) (Abcam, CA, USA) for detecting Histone

H3 protein. The ECL detection system (Millipore) was used to detect immunoreactive proteins by utilizing anti-mouse or anti-rabbit secondary antibodies (1/20,000).

Spotting test to compare the relative zeocin resistance of the transformants

Spotting test was performed to estimate the level of zeocin resistance as previously described (34). Independent transgenetic strains were cultured to reach the logarithmic phase in TAP medium and subsequently subjected to serial dilutions (1:5) with TAP medium. Aliquots (1 μL) of the diluted samples were spotted on plates supplemented with various concentrations of zeocin (0, 30, 60, and 120 mg/L) and incubated for 7–10 days under white fluorescent light (20 μmol m^{-2} s^{-1}) at 25°C.

Results and Discussion

Enhanced nuclear transgene expression in a "met1" mutant of Chlamydomonas

To date, two cytosine-specific DNA methyltransferases (i.e., *Cre10.g461750*, and *Cre12.g484600*) have been detected in the *Chlamydomonas* genome sequence (*Chlamydomonas* genomic information features v5.3.1). We isolated an insertion tagged mutant in *Cre10.g461750* tag mutant (20) (Fig. S1). The enzyme encoded by *Cre10.g461750* was expected to function at the DNA replication foci because of the presence of a protein domain (IPR022702, http://www.ebi.ac.uk/interpro/IPR022702) that functions to target this enzyme toward the replication foci (35). Thus, this gene probably encodes a maintenance-type DNA cytosine methyltransferase 1 (*Dnmt1 or Met1*), while the other one may encode a *de novo* cytosine methyltransferase. In this tagged mutant, the transcript was under detection level by semi-quantitative reverse transcription PCR analyses, whereas it was evidently detected in the mother strain CMJ030 (Fig. S2). This strain contains an additional insertion whose position has not

been identified because of the complex tag insertion and deletion detected by the sequence data of RESDA-PCR (36) products (data not shown). Therefore, the behavior of this tagged mutant is due to the disrupted *Met1* gene or unidentified gene(s). Considering this fact, we refer to this tagged mutant as the "*met1*" mutant.

The main characteristics of this tagged mutant may be different from those expected from the mutation of the maintenance type cytosine methylase; therefore, this tagged mutant has been here referred to as the "*met1*" mutant. Interestingly, the growth of this tag mutant in TAP was not different from that of the mother strain (data not shown). This suggests that the loss of DNA methylation has no significant effect on transcriptional repression required for homeostasis and effective silencing of transposons in *Chlamydomonas*.

In this study, the endogenous *SQS* gene was overexpressed in the "*met1*" mutant and in wild-type CC-124 to compare the expression levels. The linearized cDNA expression cassette was transformed by electroporation. Transformants were selected on TAP/agar plates supplemented with 15 mg/L zeocin. Considering the occurrence of decoupling through introduction, zeocin-resistant transformants were further screened by PCR for the cointegration of the *SQS* cDNA expression cassette with the marker gene *ble*. We considered these PCR-positive transformants as being successfully cotransformed (Table 1). The PCR results showed that the cotransformation ratio of *SQS* with *ble* in the "*met1*" strain was approximately 43.8% (63/144) (Table 1 and Fig. S3A), which is not significantly different from that of the wild-type CC-124 (109/210, 51.9%).

Western blotting was performed to analyze the protein accumulation levels of the PCR-positive transformants. The results showed that the western-positive ratio

detectable after 10 min exposure for the *SQS* transformants in the "*met1*" strain was 18.9% (10/53) (Table 1 and Fig. S3B), which was 4 times higher than that in the CC-124 strain (4.6%, 5/109) (Table 1). Moreover, the high protein accumulation ratio of the western-positive transformants was 30.0% (3/10) for the "*met1*" strain (Table 1 and Fig. 2A), which was approximately 1.5 times higher than that in the CC-124 strain (20.0%, 1/5) (Table 1). This suggests the ability of "*met1*" to express the *SQS* cDNA is comparable to that of the UVM strains, which have been demonstrated to bear a high potential (16, 19).

We detected 21.2 % (11/52) and 17.6% (9/51) western-positive transformants and 45.5% (5/11) and 44.4% (4/9) of highly *SQS*-expressing strains in UVM4 and UVM11, respectively (19). The main advantage of using the "*met1*" strain over UVM strains is that it possesses mating ability, which facilitates the accumulation of useful traits by genetic crossing. Thus, the "*met1*" strain could be a novel promising host cell for robust transgene expression (Fig. S4). The most probable reason for the high western-positive ratio is the successful disruption of the transcriptional silencing system caused by disabled *Dnmt1* gene, which is involved in the maintenance of DNA methylation patterns (21, 22). However, frequently detected western-negative transformants among PCR-positive transformants propose that the silencing ability has not been completely knocked out in the "*met1*" strain. This is also the case of the UVM strains (19). The retained silencing ability of the "*met1*" strain may be due to DNA methylation-independent silencing pathways (12, 13).

The above results show that this tag-inserted mutant is one of the promising hosts for plant cell factories. However, this strain bears another tag besides that at the *Met1* gene. It is essential to generate a backcrossed strain that bears only one tag at the *Met1* gene for a robust genetic background. Considering this observation, in this

study, expression analysis in "*met1*" was limited to the *Chlamydomonas SQS* gene.

Improved expression of heterologous SSL genes by the ble2A expression system in the UVM4 strain

First, we attempted to overexpress *CrSSL3* in the wild-type CC-1690 and UVM4 strains using an expression vector similar to that used for *SQS* expression, in which the *SQS* ORF was replaced with *CrSSL3* ORF and the *ble* marker gene was switched to *aadA*. However, no western-positive transgenic lines were found in the large number of PCR-positive transformants, which accounted for 261 transformants in CC-1690 and 294 in UVM4. This situation contrasts with that of *SQS* for which western-positive transformants were easily found (Table 1). Our previous success for *SQS* expression and current failure to find the western-positive *SSL3* transformants in the UVM4 strain clearly demonstrate that this strain is not a useful strain to heterologously express versatile genes of interest.

Then, we tested the *CrSSL* genes expression using the ble2A system to investigate whether it could be useful for heterologous expression of hydrocarbon production-related *Botryococcus* genes. We subcloned the *CrSSL3* or *CrSSL1* cDNA into the ble2A nuclear expression vector to generate ble2A-SSL fusion ORFs. The fused ORFs were placed under the control of the modified *Hsp70A/RbcS2* promoter (PAR4), which was revised to contain four copies of the first intron of *RbcS2* between the *Hsp70A* and *RbcS2* promoters (Fig. 1B). Then, the linearized plasmids were introduced into the wild-type CC-1690 and UVM4 strains. The cotransformation- and western-positive ratios for the *SSL3* in the UVM4 strain were approximately 51.3% (74/144) and 10.4% (5/48), respectively, which were 1.5 times and 5.0 times higher

than those in the CC-1690 strain [34.0% (49/144) and 2.0% (1/49)] (Table 2, Fig. S5). Moreover, the protein accumulation ratio in the western-positive transformant was 40.0% (2/5) in the UVM4 strain, whereas no *SSL3* strongly expressing transformants (0/1) were found in the CC-1690 strain (Table 2 and Fig. 2B). This intimates the high ability of the ble2A system for heterologous expression of cDNAs.

In *ble2A-CrSSL3* transformants of UVM4, the fused products were specifically detected in highly SSL3-accumulating transformants, whereas in *ble2A-CrSSL1* transformants of C-9 (the wild-type), the fused products were detected even in weakly *CrSSL1*-expressing transformants (Fig. 2B, Fig. S6B). These results suggest that ribosome skipping efficiency at the ble2A coding region is strongly affected by the following ORF or by the different characteristics of the strains used. Moreover, we found that the ribosome-skipping efficiency of the *ble2A-SSL3* was not uniform even among the transformants (Fig. 2B). Transformant 3 is one of the highly SSL3-expressing transformants; in this transformant, the fused protein level was almost identical to that of the processed SSL3, whereas in transformant 5, which also highly expresses *CrSSL3*, the level of the fused product prominently exceeded that of the processed protein. Interestingly, in weakly expressing transformants, no fused protein was detected. The sequence of FMDV 2A induces ribosome-skipping during translation to generate discrete products from single ORFs (17, 37, 38), in which the translational release factors, eRF1 and eRF3, play an important role (39). However, the relation between the expression level of the protein and the ribosome-skipping ratios remains unknown.

Through western analyses of the transformants, we showed that SSL3-positive transformants were much more frequently detected in the UVM4 strain (5/48) than in the wild-type strain CC-1690 (1/49). This result is in agreement with our previous

experiments showing the efficient expression of the *SQS* cDNA from *Chlamydomonas* in the UVM4 and UVM11 strains (19). Therefore, UVM strains are much better hosts than wild-type strains, albeit not ideal, for the expression of various types of heterologous cDNAs.

In our previous experiment in which the *Hsp70A/RbcS2* promoter was used to stimulate the above-mentioned *SQS* cDNA, the expression levels varied prominently among transformants even in the UVM strains. Therefore, in this study, we used a modified *Hsp70A/RbcS2* promoter (PAR4), which contained four introns to enhance the transcriptional ability. However, prominent variation of transgene expression levels was still observed as in the case of the non-modified *Hsp70A/RbcS2* promoter (Fig. 2B, Fig. S6B), and the effects between the two promoters were not significantly different. This suggests that the PAR4 promoter is still not sufficiently strong to overcome the remaining variation of transgene expression levels in the UVM4 strain, regardless of its apparently alleviated silencing ability.

Therefore, the expression of a gene of interest using the ble2A system in the UVM or "*met1*" strains surely enhances the possibility to find transformants that highly accumulate the target protein. This combinational method may be useful for heterologous expression of almost all transgenes.

Comparative analysis of the ble2A-CrSSL1 mRNA levels in the western-positive and -negative C-9 transformants.

To compare the ble2A-CrSSL1 mRNA levels, total mRNA was isolated from C-9 (wild-type) and transformants of ble2A-CrSSL1 that strongly (SSL1-5 and -34) and weakly (SSL-7 and -39) expressed CrSSL1 (Fig. S7). The results of RT-PCR clearly

showed that the relative abundance of ble2A-CrSSL1 mRNA was higher in SSL1-5 and -34 than in SSL-7 and -39 (Fig. 3). Thus, the data shows that the differences in the SSL1 protein levels are attributable mainly to variation in the mRNA levels. *Chlamydomonas* has no RNA-dependent RNA polymerase gene (40) that is essential to exert strong RNA interference (41). Therefore, the observed variation in transgene expression might be mainly due to position-dependent levels of transcriptional gene silencing (TGS) but not due to post-transcriptional gene silencing (PTGS), which has been suggested to be the main factor in position-dependent differences in single transgene expression in *Arabidopsis* (42, 43).

Determination of the relationship between target protein expression levels and zeocin resistance

The mechanism of robust expression of recombinant proteins by transcriptionally fusing their genes to the *ble* marker gene is not yet well characterized. One of the probable reasons is that *ble* functions by sequestration of the antibiotic through one-to-one stoichiometric binding, but not through enzymatic inactivation of zeocin (18). Therefore, presumably high levels of *ble* expression are required for survival.

Ble and SSL proteins must be produced in 1:1 ratio under when mRNA is fully translated without breaks and ribosome-skipping occurs, as expected, perfectly at the 2A region. In turn, this leads to the simple speculation that zeocin resistance and amount of SSL should show a positive relationship. If so, screening of high zeocin-resistant transformants could be a practical strategy to find highly expressed target gene. Based on this idea, the relationship between zeocin-resistance and SSL levels was analyzed. We measured the levels of zeocin resistance in SSL western-positive

transformants by spotting transformant cells on agar plates containing various zeocin concentrations.

First, we confirmed that all the transformants chosen for analysis showed no significant growth differences on non-drug-containing TAP plates (Fig. 2C). UVM4 transformants 3 and 5 (UVM4-3 and UVM4-5, respectively) were expected to show relatively rapid growth on plates containing zeocin at high concentration, because they accumulated relatively high levels of SSL3 protein compared to other transformants. As expected, UVM4-3 showed the best growth on the plate containing 120 mg/L zeocin, whereas the growth of UVM4-5 was slightly but significantly slower than that of UVM4-1 on the plates containing 30 mg/L or 60 mg/L zeocin, of which the SSL3 expression level was prominently lower than that of UVM4-5 (Fig. 2). Moreover, no transformants survived on the plate containing 200 mg/L zeocin. These results show that the correlation between the target protein expression levels and the zeocin-resistance levels is not so strict in the transformants obtained using the ble2A expression system. A similar result was observed in *CrSSL1* expressing C-9 wild-type transformants (Fig. S6). The transformants accumulating SSL1 protein at low levels (e.g., SSL1-39) showed growth rate not apparently different from that of the highly expressing transformants (e.g., SSL1-34) at the various zeocin concentrations (Fig. S6). Consistent with the results of these analyses, we failed to sort highly target protein-expressing transformants of *CrSSL1* among the transformants survived on the plate containing 120 mg/L zeocin, irrespective of the drastically decreased number of transformants on the plates (data not shown). Therefore, zeocin-resistance screening does not appear to be a practical method to identify the transformants highly expressing target proteins.

The above is also the case for the two independent expression cassette system,

i.e., target gene expression cassette and marker gene expression cassette. Using the expression plasmid containing the independent cassettes of *CrSSL1* and *ble* in vicinity, we could not obtain the *SSL1*-highly expressing lines among the transformants appeared on the plate containing 120 mg/L of zeocin (data not shown).

We suspect that the *ble* sequence may have an unknown effect leading to enhancement of translation or prevention of translational stall, and it appears to work efficiently only when the sequence is directly ligated to the target ORF. Very efficient expression of *ble-CrGFP,* in which *ble* is directly connected with *CrGFP,* is another example of a system showing similar ble-fusion effect (9).

Further studies are essential to unveil the mechanism of action of the ble2A system and to explain why it works efficiently in overcoming the silencing of heterologous transgenes. The combinational use of the ble2A nuclear expression vector and UVM4 or "*met1*" strains for foreign gene expression could enhance the utility of *Chlamydomonas* as plant cell factory for producing biofuel, biopharmaceuticals, and other valuable compounds.

Acknowledgments

The authors would like to thank Dr. R. Bock (MPI-MP, Germany) for kindly providing the *C. reinhardtii* UVM strains; and Ru Zhang, Weronika Patena, Spencer Gang, Sean Blum, Rebecca Yue, Arthur Grossman, Martin Jonikas (Carnegie Institution for Science, Stanford, CA, USA) for the *met1* mutant. This research was supported by Japan Science and Technology Agency (JST), Core Research for Evolutional Science and Technology (CREST), and the U.S. National Science Foundation grant MCB-1146621.

References

1. **Pulz, O., and Gross, W.:** Valuable products from biotechnology of microalgae. Appl. Microbiol. Biotechnol., 65, 635–648 (2004).

2. **Rasala, B., Lee, P., Shen, Z., Briggs, S., Mendez, M., and Mayfield, S.:** Robust expression and secretion of *Xylanase1* in *Chlamydomonas reinhardtii* by fusion to a selection gene and processing with the FMDV 2A peptide. PLoS One, 7, e43349 (2012).

3. **Spolaore, P., Joannis-Cassan, C., Duran, E., and Isambert, A.:** Commercial applications of microalgae. J. Biosci. Bioeng., 101, 87–96 (2006).

4. **Harris, E.:** *Chlamydomonas* as a Model Organism. Annu. Rev. Plant Physiol. Plant Mol. Biol., 52, 363–406 (2001).

5. **Hippler, M., Redding, K., and Rochaix, J.:** *Chlamydomonas* genetics, a tool for the study of bioenergetic pathways. Biochim. Biophys. Acta, 1367, 1–62 (1998).

6. **Pedersen, L., Geimer, S., and Rosenbaum, J.:** Dissecting the molecular mechanisms of intraflagellar transport in *Chlamydomonas*. Curr. Biol., 16, 450–459 (2006).

7. **Schmidt, M., Gessner, G., Luff, M., Heiland, I., Wagner, V., Kaminski, M., Geimer, S., Eitzinger, N., Reissenweber, T., Voytsekh, O., and other 3 authors:** Proteomic analysis of the eyespot of *Chlamydomonas reinhardtii* provides novel insights into its components and tactic movements. Plant Cell, 18, 1908–1930 (2006).

8. **Merchant, S., Prochnik, S., Vallon, O., Harris, E., Karpowicz, S., Witman, G., Terry, A., Salamov, A., Fritz-Laylin, L., Marechal-Drouard, L., and other 107 authors:** The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. Science, 318, 245–250 (2007).

9. **Fuhrmann, M., Oertel, W., and Hegemann, P.:** A synthetic gene coding for the green fluorescent protein (GFP) is a versatile reporter in *Chlamydomonas reinhardtii*. Plant J., 19, 353–361 (1999).

10. **Shao, N., and Bock, R.:** A codon-optimized luciferase from *Gaussia princeps* facilitates the in vivo monitoring of gene expression in the model alga *Chlamydomonas reinhardtii.* Curr. Genet., 53, 381–388 (2008).

11. **Tam, L., and Lefebvre, P.:** Cloning of flagellar genes in *Chlamydomonas reinhardtii* by DNA insertional mutagenesis. Genetics, 135, 375–384 (1993).

12. **Cerutti, H., Johnson, A., Gillham, N., and Boynton, J.:** Epigenetic silencing of a foreign gene in nuclear transformants of *Chlamydomonas*. Plant Cell, 9, 925–945 (1997).

13. **Schroda, M.:** RNA silencing in *Chlamydomonas*: mechanisms and tools. Curr. Genet., 49, 69–84 (2006).

14. **Eichler-Stahlberg, A., Weisheit, W., Ruecker, O., and Heitzer, M.:** Strategies to facilitate transgene expression in *Chlamydomonas reinhardtii*. Planta, 229, 873–883 (2009).

15. **Schroda, M., Blocker, D., and Beck, C.:** The HSP70A promoter as a tool for the improved expression of transgenes in *Chlamydomonas*. Plant J., 21, 121–131 (2000).

16. **Neupert, J., Karcher, D., and Bock, R.:** Generation of *Chlamydomonas* strains that efficiently express nuclear transgenes. Plant J., 57, 1140–1150 (2009).

17. **de Felipe, P., Luke, G., Hughes, L., Gani, D., Halpin, C., and Ryan, M.:** E unum pluribus: multiple proteins from a self-processing polyprotein. Trends Biotechnol., 24, 68–75 (2006).

18. **Rasala, B., Barrera, D., Ng, J., Plucinak, T., Rosenberg, J., Weeks, D., Oyler, G., Peterson, T., Haerizadeh, F., and Mayfield, S.:** Expanding the spectral palette

of fluorescent proteins for the green microalga *Chlamydomonas reinhardtii*. Plant J., 74, 545–556 (2013).

19. **Kong, F., Yamasaki, T., and Ohama, T.:** Expression levels of domestic cDNA cassettes integrated in the nuclear genomes of various *Chlamydomonas reinhardtii* strains. J. Biosci. Bioeng., 117, 613–616 (2014).

20. **Zhang, R., Patena, W., Armbruster, U., Gang, S., Blum, S., and Jonikas, M.:** High-Throughput Genotyping of Green Algal Mutants Reveals Random Distribution of Mutagenic Insertion Sites and Endonucleolytic Cleavage of Transforming DNA. Plant Cell, 26, 1398–1409 (2014).

21. **Bird, A.:** DNA methylation patterns and epigenetic memory. Genes Dev., 16, 6–21 (2002) .

22. **Fuks, F., Hurd, P., Wolf, D., Nan, X., Bird, A., and Kouzarides, T.:** The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. J. Biol. Chem., 278, 4035–4040 (2003).

23. **Niehaus, T., Okada, S., Devarenne, T., Watt, D., Sviripa, V., and Chappell, J.:** Identification of unique mechanisms for triterpene biosynthesis in *Botryococcus braunii*. Proc. Natl. Acad. Sci. USA, 108, 12260–12265 (2011).

24. **Meslet-Cladiere, L., and Vallon, O.:** Novel shuttle markers for nuclear transformation of the green alga *Chlamydomonas reinhardtii*. Eukaryot. Cell, 10, 1670–1678 (2011).

25. **Ruecker, O., Zillner, K., Groebner-Ferreira, R., and Heitzer, M.:** *Gaussia*-luciferase as a sensitive reporter gene for monitoring promoter activity in the nucleus of the green alga *Chlamydomonas reinhardtii*. Mol. Genet. Genomics, 280, 153–162 (2008).

26. **Urban, A., Neukirchen, S., and Jaeger, K.:** A rapid and efficient method for site-directed mutagenesis using one-step overlap extension PCR. Nucleic Acids Res., 25, 2227–2228 (1997).

27. **Zhou, J., and Blissard, G.:** Identification of a GP64 subdomain involved in receptor binding by budded virions of the baculovirus *Autographica californica* multicapsid nucleopolyhedrovirus. J. Virol., 82, 4449–4460 (2008).

28. **Sanger, F., Nicklen, S., and Coulson, A.:** DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA, 74, 5463–5467 (1977).

29. **Gorman, D., and Levine, R.:** Cytochrome f and plastocyanin: their sequence in the photosynthetic electron transport chain of *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. USA, 54, 1665–1669 (1965).

30. **Shimogawara, K., Fujiwara, S., Grossman, A., and Usuda, H.:** High-efficiency transformation of *Chlamydomonas reinhardtii* by electroporation. Genetics, 148, 1821–1828 (1998).

31. **Cao, M., Fu, Y., Guo, Y., and Pan, J.:** *Chlamydomonas* (Chlorophyceae) colony PCR. Protoplasma, 235, 107–110 (2009).

32. **Rasala, B., Muto, M., Lee, P., Jager, M., Cardoso, R., Behnke, C., Kirk, P., Hokanson, C., Crea, R., Mendez, M., and Mayfield, S.:** Production of therapeutic proteins in algae, analysis of expression of seven human proteins in the chloroplast of *Chlamydomonas reinhardtii*. Plant Biotechnol. J., 8, 719–733 (2010).

33. **Yamasaki, T., Voshall, A., Kim, E., Moriyama, E., Cerutti, H., and Ohama, T.:** Complementarity to an miRNA seed region is sufficient to induce moderate repression of a target transcript in the unicellular green alga *Chlamydomonas reinhardtii*. Plant J., 76, 1045–1056 (2013).

34. **Yamasaki, T., Miyasaka, H., and Ohama, T.:** Unstable RNAi effects through epigenetic silencing of an inverted repeat transgene in *Chlamydomonas reinhardtii*. Genetics, 180, 1927–1944 (2008).

35. **Rountree, R., Bachman, E., and Baylin, B.:** DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. Nat. Genet., 25, 269–277 (2000).

36. **Gonzalez-Ballester, D., de Montaigu, A., Galvan, A., and Fernandez, E.:** Restriction enzyme site-directed amplification PCR: a tool to identify regions flanking a marker DNA, Anal. Biochem., 340, 330–335 (2005).

37. **de Felipe, P.:** Skipping the co-expression problem: the new 2A "CHYSEL" technology. Genet. Vaccines Ther., 2, 13 (2004).

38. **Donnelly, M., Luke, G., Mehrotra, A., Li, X., Hughes, L., Gani, D., and Ryan, M.:** Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. J. Gen. Virol., 82, 1013–1025 (2001).

39. **Doronina, V., Wu, C., de Felipe, P., Sachs, M., Ryan, M., and Brown, J.:** Site-specific release of nascent chains from ribosomes at a sense codon. Mol. Cell Biol., 28, 4227–4239 (2008).

40. **Cerutti, H., and Casas-Mollano, A**.: On the origin and functions of RNA-mediated silencing from protists to man. Curr. Genet., 50, 81–99 (2006).

41. **Yamasaki, T., Miyasaka, H., and Ohama, T**.: Unstable RNAi effect through methyl-CpG related transcriptional repression of inverted repeat transgene in *Chlamydomonas reinhardtii*. Genetics, 180, 1927–1944 (2008).

42. **Butaye, M., Goderis, J., Wouters, F., Pues, M., Delaure, L., Broekaert. F., Depicker, A., Cammue, P., and de Bolle, F**.: Stable high-level transgene expression

in *Arabidopsis thaliana* using gene silencing mutants and matrix attachment regions. Plant J., 39, 440–449 (2004).

43. **Schubert, D., Lechtenberg, B., Forsbach, A., Gils, M., Bahadur, S., and Schmidt, R.:** Silencing in *Arabidopsis* T-DNA transformants: the predominant role of a gene-specific RNA sensing mechanism versus position effects. Plant Cell, 16, 2561–2572 (2004).

Figure legends

Figure 1. Schematic representation of the transformation vectors. (A) Expression cassette for *Chlamydomonas SQS* cDNA and *ble* marker gene. in, first intron of the *Chlamydomonas RbcS2* gene; *RbcS2* T, *Chlamydomonas RbcS2* terminator; gp-64-tag, gp-64 gene sequence for epitope tag peptide. The arrows show the location of the PCR primers used for the cotransformation assay. (B) Expression cassette for the codon-adjusted *Botryococcus braunii SSL3* cDNA. *PAR4, Hsp70A/RbcS2* promoter modified to contain four copies of the first intron of *RbcS2*; 2A, FMDV 2A peptide, *CrSSL3*, codon-optimized *SSL3* cDNA. The amino acid sequence of 2A is denoted by an asterisk to indicate the cleavage site.

Figure 2. Expression levels of SQS and SSL3-positive transformants analyzed by western blotting. (A) The expression levels of SQS-positive transformants in the "*met1*" strain are shown in the upper gel, while the lower gel shows the result of histone H3 in which equal amounts of proteins were loaded. (B) The expression of SSL3-positive transformants in the CC-1690 and UVM4 strains (1 through 5) is shown. The amount of the proteins in the bands was quantified using the Image J software (http://imagej.nih.gov/ij/). The band intensities are shown as relative

intensity units to the positive controls, which are one of the SQS transformants in the UVM4 strain (19) for (A) and histone H3 for (B). Signal intensity values above 1.0 for SQS and above 0.3 for SSL3 were attributed to strongly expressing transformants and marked with asterisks. The arrowhead indicates unprocessed fusion protein. (C) Zeocin resistance spotting test for SSL3 western-positive UVM4 transformants. Spotted cell numbers are shown below the spots.

Fig. 3. Comparative analysis of the ble2A-CrSSL1 mRNA levels by semi-quantitative reverse transcription PCR. (A) RT-PCR results at the end of 30 cycles using cDNA prepared from SSL1-7, -5. -34, -39, and C-9 (untransformed wild-type) for partial amplification. Asterisks indicate the strains with high expression levels of CrSSL1 protein (see Fig. S5). +RT and −RT denote the reactions with and without reverse transcriptase, respectively. (B) RT-PCR results at the end of 20 cycles, which show that equal amounts of mRNA were used. An endogenous *CBLP* cDNA was amplified. *CBLP, Chlamydomonas β* subunit-like polypeptide. See Fig. S6 for details.

Table 1. Expression of *SQS* cDNA cassette in various *C. reinhardtii* strains.

Host strain	Cotransformation ratio[a]	Western blot-positive ratio[b]	Ratio of highly expressing transformant[c] in Western blot-positive transformants
CC-124[e] (WT[d])	109/210 (51.9%)	5/109 (4.6%)	1/5 (20.0%)
UVM4[e]	52/87 (59.8%)	11/52 (21.2%)	5/11 (45.5%)
UVM11[e]	51/87 (58.6%)	9/51 (17.6%)	4/9 (44.4%)
"*met1*"	63/144 (43.8%)	10/53 (18.9%)	3/10 (30.0%)

[a]Number of PCR-positive transformants/number of transformants analyzed.

[b]Number of Western blot-positive transformants/number of PCR-positive transformants.

[c]Number of high-protein accumulation transformants/number of Western blot-positive transformants.

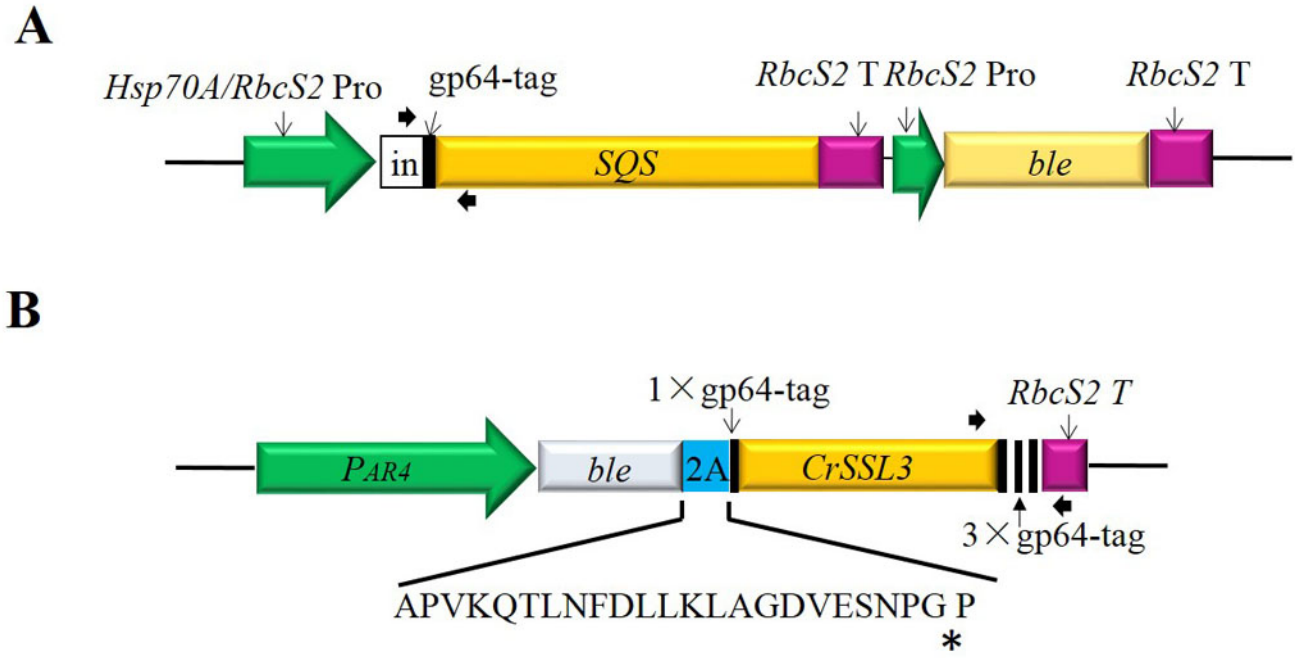
[d]Wild-type.

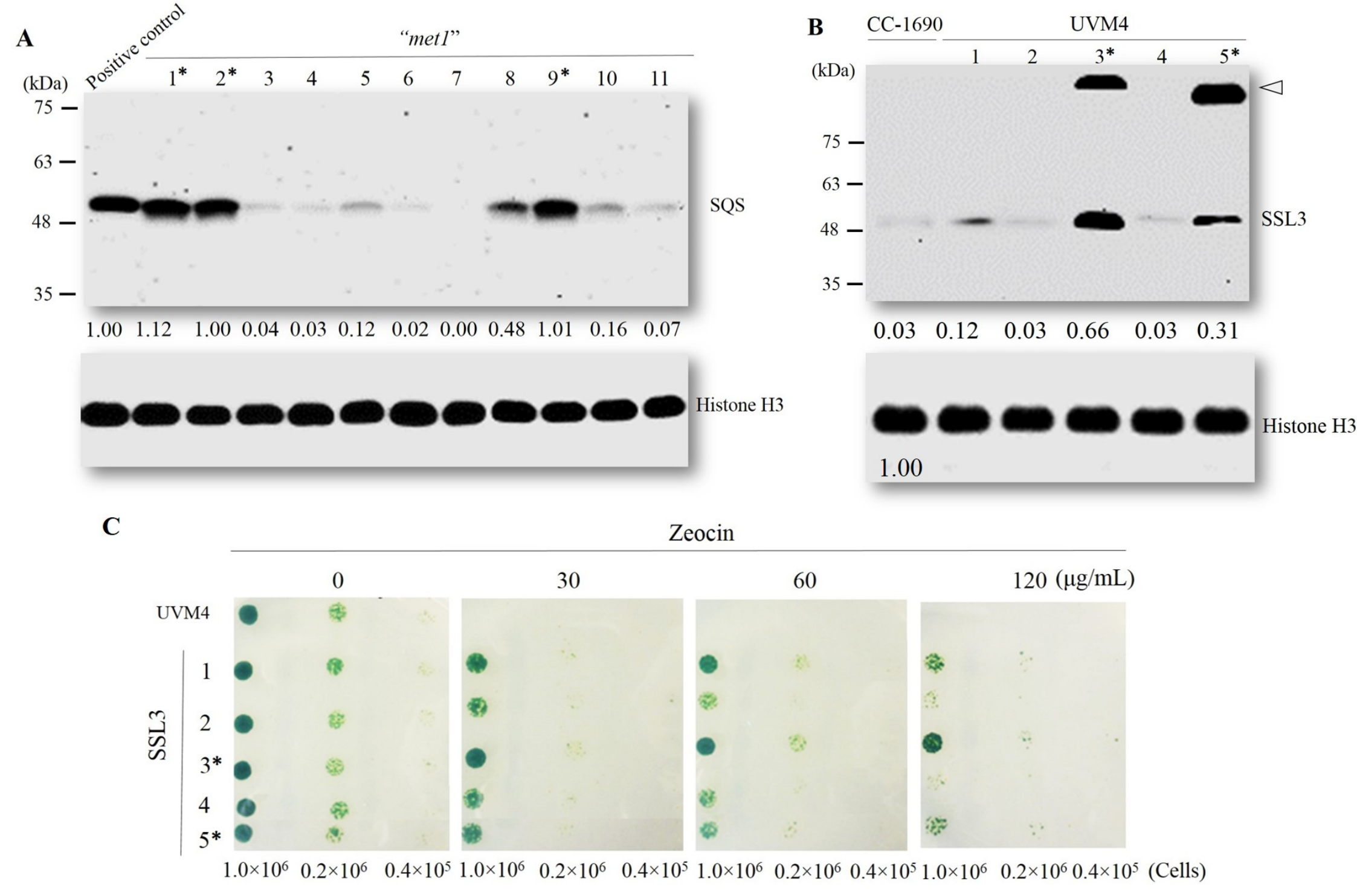
[e]Data were taken from our previous experiment in Kong et al. (19).

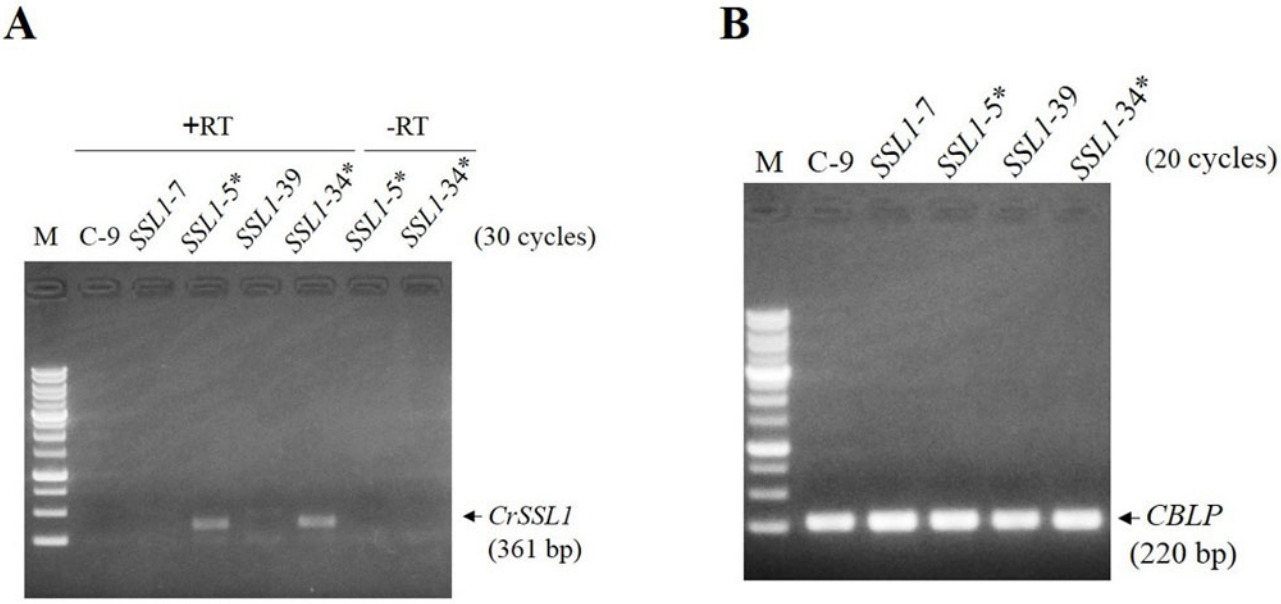
Table 2. Expression of SSL3 cDNA using the ble2A system in CC-1690 and UVM4 strains.

Host strain	Cotransformation ratio[a]	Western blot-positive ratio[b]	Ratio of highly expressing transformant[c] in Western blot-positive transformants
CC-1690 (WT[d])	49/144 (34.0%)	1/49 (2.0%)	0/1 (0.0%)
UVM4	74/144 (51.3%)	5/48 (10.4%)	2/5(40.0%)

See the footnotes of Table1 for a, b, c, and d.


A
Hsp70A/RbcS2 Pro
gp64-tag
RbcS2 T
RbcS2 Pro
RbcS2 T
in
SQS
ble
B
1×gp64-tag
RbcS2 T
PAR4
ble
2A
CrSSL3
3×gp64-tag
APVKQTLNFDLLKLAGDVESNPG P
*


A
Positive control
"met1"
(kDa)
1* 2* 3 4 5 6 7 8 9* 10 11
75
63
48
35
SQS
1.00 1.12 1.00 0.04 0.03 0.12 0.02 0.00 0.48 1.01 0.16 0.07
Histone H3
B
CC-1690
UVM4
(kDa)
1 2 3* 4 5*
75
63
48
35
SSL3
0.03 0.12 0.03 0.66 0.03 0.31
Histone H3
1.00
C
Zeocin
0 30 60 120 (μg/mL)
UVM4
SSL3
1
2
3*
4
5*
1.0×10⁶ 0.2×10⁶ 0.4×10⁵ 1.0×10⁶ 0.2×10⁶ 0.4×10⁵ 1.0×10⁶ 0.2×10⁶ 0.4×10⁵ 1.0×10⁶ 0.2×10⁶ 0.4×10⁵ (Cells)


A
+RT
-RT
M
C-9
SSL1-7
SSL1-5*
SSL1-39
SSL1-34*
SSL1-5*
SSL1-34*
(30 cycles)
CrSSL1
(361 bp)
B
M
C-9
SSL1-7
SSL1-5*
SSL1-39
SSL1-34*
(20 cycles)
CBLP
(220 bp)