Plant Cell Rep (2011) 30:113–124
DOI 10.1007/s00299-010-0949-y

ORIGINAL PAPER

High-level accumulation of recombinant miraculin protein in transgenic tomatoes expressing a synthetic *miraculin* gene with optimized codon usage terminated by the native miraculin terminator

Kyoko Hiwasa-Tanase · Mpanja Nyarubona · Tadayoshi Hirai · Kazuhisa Kato · Takanari Ichikawa · Hiroshi Ezura

Received: 3 August 2010 / Revised: 27 October 2010 / Accepted: 29 October 2010 / Published online: 13 November 2010

Abstract In our previous study, a transgenic tomato line that expressed the *MIR* gene under control of the *cauliflower mosaic virus 35S* promoter and the *nopaline synthase* terminator (*tNOS*) produced the taste-modifying protein miraculin (MIR). However, the concentration of MIR in the tomatoes was lower than that in the *MIR* gene's native miracle fruit. To increase MIR production, the native *MIR* terminator (*tMIR*) was used and a synthetic gene encoding MIR protein (*sMIR*) was designed to optimize its codon usage for tomato. Four different combinations of these genes and terminators (*MIR-tNOS, MIR-tMIR, sMIR-tNOS* and *sMIR-tMIR*) were constructed and used for transformation. The average MIR concentrations in *MIR-tNOS*, *MIR-tMIR*, *sMIR-tNOS* and *sMIR-tMIR* fruits were 131, 197, 128 and 287 μg/g fresh weight, respectively. The MIR concentrations using *tMIR* were higher than those using *tNOS*. The highest MIR accumulation was detected in *sMIR-tMIR* fruits. On the other hand, the MIR concentration was largely unaffected by *sMIR-tNOS*. The expression levels of both *MIR* and *sMIR* mRNAs terminated by *tMIR* tended to be higher than those terminated by *tNOS*. Read-through mRNA transcripts terminated by *tNOS* were much longer than those terminated by *tMIR*. These results suggest that *tMIR* enhances mRNA expression and permits the multiplier effect of optimized codon usage.

Keywords Miraculin · Codon optimization · Miraculin terminator · Transgenic tomato · Read-through

Abbreviations

GUS	β-Glucronidase
MIR	Miraculin
sMIR	Synthesized MIR
NOS	Nopaline synthase

Introduction

Plants provide many advantages for the production of valuable heterologous proteins over other production systems in terms of practicality, economy and safety (Twyman et al. 2003; Desai et al. 2010). In fact, the production and storage costs with plant systems are low compared to other systems such as mammalian cell culture and microbial fermentation (Desai et al. 2010). Plants are also a convenient system for large-scale production and have a lower risk of contamination by human pathogenic microorganisms (Giddings et al. 2000; Desai et al. 2010). Additionally, when target proteins are produced in the edible part of a plant, the protein can be consumed raw as an edible vaccine (Mason et al. 2002). However, the use of plants as expression hosts has several constraints. The primary limitation is the low level of protein accumulation (Daniell et al. 2001). Ways of improving transcription and

Communicated by K. Toriyama.

K. Hiwasa-Tanase · M. Nyarubona · T. Hirai · K. Kato · H. Ezura (✉)
Graduate School of Life and Environmental Sciences, University of Tsukuba, Tennodai 1-1-1, Tsukuba, Ibaraki 305-8572, Japan
e-mail: ezura@gene.tsukuba.ac.jp

T. Ichikawa
Scientific Senior Manager of Technology Center, Team Leader of Accreditation and Evaluation Team for the New University, Okinawa Institute of Science and Technology Promotion Corporation, 1919-1 Tancha, Onna-son, Okinawa 904-0411, Japan

translation levels include testing various promoters and terminators, and optimizing codon usage.

The compatibility of the promoter with the host plant is crucial for high expression of the target gene (Twyman et al. 2003; Desai et al. 2010). However, the 3′-untranslated region (UTR) and terminator of the mRNA also influence the expression level of the target gene by controlling RNA transcript termination and polyadenylation (Proudfoot 2004; Gilmartin 2005), and the efficiency of transcription is different with different types of terminators (Ingelbrecht et al. 1989; Nagaya et al. 2010). In fact, the expression level can change 60-fold depending on the terminator sequence used in a transient expression assay (Ingelbrecht et al. 1989).

The preferred codon usage varies significantly among different plant species (Murray et al. 1989). Therefore, when rare codons for a host plant are used in an introduced gene, the codon becomes the limiting factor in the translation process. This restriction is especially strong in the case of molecular farming in plants, because the aim is mass production of the target protein. One of the strategies to increase translation efficiency is to modify the codons from the original sequence to more suitable ones in the host plant without changing the amino acid sequence (Gustafsson et al. 2004). By codon optimization, expression of the insect control protein gene *cryIA(b)* from *Bacillus thuringienisis* was increased up to 100-fold in transgenic tomato and tobacco (Perlak et al. 1991). Moreover, codon modification of the reductase gene (P450) from wheat increased its protein accumulation level in transgenic tobacco, demonstrating that preferred codon usage is different among plant species (Batard et al. 2000).

Miraculin (MIR) is a glycoprotein in miracle fruit (*Richadella dulcifica*), a shrub originally from West Africa (Theerasilp and Kurihara 1988). It has the unique property of changing a sour taste into a sweet taste, although MIR itself is not sweet. It is possible to provide a safe yet appetizing diet for diabetic and dieting people who require a restricted diet by utilizing the taste-modifying behavior of MIR as an alternative low-calorie sweetener. Most importantly, this property of MIR can make dieting more appealing. However, despite its great potential, miracle fruit production is limited because it is a tropical plant. Efforts to produce recombinant MIR have succeeded using plants such as tomato, lettuce and strawberry as a host (Sun et al. 2006a, 2007; Sugaya et al. 2008). Among these species, tomato was the most suitable host for MIR production (Yano et al. 2010). The level of MIR accumulation remained steady over multiple generations, and the introduced *MIR* gene was stably inherited (Yano et al. 2010). In the studies, the *MIR* gene was driven by a cauliflower mosaic virus (*CaMV*) *35S* promoter and terminated by the *nopaline synthase* (*NOS*) terminator, and the MIR content was <1% of total soluble protein in the transgenic tomato fruit (Sun et al. 2007). In contrast, the MIR content in miracle fruit is around 10% of total soluble protein (Theerasilp and Kurihara 1988). It means that recombinant MIR protein is produced in miracle fruit more efficiently. Therefore, to improve transcript efficiency, we isolated the *MIR* terminator and used it instead of the *NOS* terminator to produce MIR in transgenic tomatoes. In addition, the codon usage of the *MIR* gene was optimized to reflect frequently used codons in the tomato, and the effect on the translation process was evaluated.

Materials and methods

Isolation of the *MIR* terminator

Aliquots (4 μg) of genomic DNA isolated from miracle fruit leaves as described by Rogers and Bendich (1985) were digested with the restriction enzymes *Eco*RV, *Pvu*II and *Sca*I, ligated with the specific sequence adapter using Ligation high (TOYOBO, Osaka, Japan) and then used as a template for polymerase chain reaction (PCR) amplification. To determine the *MIR* terminator sequences, PCR was performed using LA Taq (Takara-Bio Inc., Otsu, Japan) with an adapter-specific primer (AP1) and MIR-specific primers (MIR1-1, MIR1-2), and then nested PCR was performed with the adapter-specific primer (AP2) and an MIR-specific primer (MIR2). The PCR products were ligated into pGEM®-T Easy vector (Promega, Madison, WI, USA) and sequenced. The primer sequences were adapter, 5′-GTAATACGACTCACTATAGGGCACGCG TGGTCGACGGCCCGGGCTGGT-3′; AP1, 5′-CCATC CTAATACGACTCACTATAGGGC-3′; AP2, 5′-CTATA GGGCACGCGTGGT-3′; MIR1-1, 5′-ACAACTCTGGGT GGACAAACGAAGCTGCCGTT-3′; MIR1-2, 5′-GGAGT TTCTCTCCGTCTATGTCAAGAACCGGATTG-3′; and MIR2, 5′-GCCGAATCCGCTGCACTAAGCAGTGGT TT-3′.

Characterization of the *MIR* terminator

To assay the termination efficiency of different length *MIR* terminators, a transient assay was performed. *MIR* terminator fragments of 147, 278, 508 and 1,085 bp were amplified by PCR using an added forward primer *Sac*I restriction site and an added reverse primer *Eco*RI restriction site (Table 1, Fig. 1). Each *MIR* terminator fragment was used to replace the *NOS* terminator fragment in plasmid pBI121 by insertion between its *Sac*I and *Eco*RI restriction sites. The resulting plasmids, named *t147, t278, t508* and *t1085*, contained the *β-glucuronidase* (*GUS*) gene flanked by the *MIR* terminator and under control of the *CaMV 35S* promoter. Each plasmid and pBI121 as a control was transferred to *Agrobacterium*

Table 1 Sequences of oligonucleotides used in the study

Name	Sequence (5′ to 3′)
tMIR-start	**TCTAAGGAGCTC**TTGGGTTTGGGGGTGGTTTTTCCA
tMIR-146R	**GCCAGTGAATTC**TCGTACACGTCAGAAACACAACGCT
tMIR-287R	**GCCAGTGAATTC**ACACTCTACTTGTGCTTTCTTGCAC
tMIR-508R	**GCCAGTGAATTC**CTACAACGTTACGAAACGTTCCTTAA
tMIR-1085R	**GCCAGTGAATTC**GTCGCTGAATAAAGGTTAGTATTGA
MIR-start	CACCCAATCCGGTTCTTGAC
MIR-stop	TTAGAAGTATACGGTTTTGTTGAACTCGAATG
sMIR-start	GAGCTTACGATGCTTTCTCTTAGC
sMIR-stop	TATCAGAAAGTGCCAATCGACGC
tNOS-R	TCCTAGTTTGCGCGCTATATTT
tMIR-R	CGTTCCTTAATGTGTGTTCAAG
LB-R1	ATTCAGGCTGCGCAACTG
LB-R2	GGTGCCGTAAAGCACTAAATC

Restriction enzyme recognition sites in sequences are shown in bold font

tumefaciens GV2260 (Deblaere et al. 1985) using the method of Shen and Forde (1989).

Sterilized tomato seeds (*Solanum lycopersicum* cv. Micro-Tom) were sown on Murashige and Skoog's medium (1962) and cultured at 25°C with 16 h of light per day for 10 days. Cotyledons of the seedlings were inoculated with plasmid-containing *Agrobacterium* and co-cultivated according to Sun et al. (2006b). After co-cultivation for 3 days, inoculated cotyledons were washed with sterilized water to remove the *Agrobacterium* and a GUS assay was performed according to the method of Jefferson et al. (1987) to evaluate the influence of the *MIR* terminators of different lengths on expression efficiency.

Codon modification of the *MIR* gene

The codon usage table for tomato (http://www.kazusa.or.jp/codon/index.html) was used to eliminate rare codons (less than 10%) in the *MIR* gene. ATTTA sequences, which are known to destabilize transcripts (Gutiérrez et al. 1999), were also removed. Additionally, mRNA secondary structure formation was minimized by original software of Invitrogen. The optimized *MIR* gene was synthesized by outsourcing (Invitrogen, Tokyo, Japan) and was cloned into pUCminusMCS. When the optimized *MIR* gene was synthesized, *Xba*I and *Sac*I restriction sites were added to the 5′ and 3′ ends of the gene, respectively. Native *MIR* and synthetic *sMIR* share 73% identity at the nucleotide sequence level (Fig. 2).

Construction of plasmids and transformation into tomato

To evaluate the individual influences of the *MIR* terminator and the codon-modified *MIR* gene on MIR accumulation in tomato, we prepared four different constructs for introduction into tomato by combining either *MIR* or *sMIR* with either the 508 bp *MIR* terminator (*tMIR*) or the *NOS* terminator (*tNOS*): *MIR-tNOS, MIR-tMIR, sMIR-tNOS* and *sMIR-tMIR*.

The native *MIR* gene was isolated from genomic DNA of miracle fruit leaves based on the published DNA sequence of *MIR* (GenBank accession number AB512278). A pair of specific primers (forward primer: 5′-TTTTTTCTAGA ATGAAGGAATTAACAATGCT-3′, and reverse primer: 5′-TTTTTTGAGCTCTTAGAAGTATACGGTTTTGT-3′) containing *Xba*I and *Sac*I sites, respectively, was used to amplify the coding region of *MIR*. The amplification reaction using KOD-Plus (TOYOBO, Osaka, Japan) consisted of 94°C for 3 min, 35 cycles of amplification (94°C for 15 s, 56°C for 25 s and 68°C for 50 s) and a final extension at 68°C for 10 min. The PCR product was cloned into the *Xba*I and *Sac*I sites of pBI121 after removing the *GUS* coding region via the restriction enzyme sites. The *sMIR* gene was subcloned from pUCminusMCS into pBI121 using the *Xba*I and *Sac*I sites in the same way as with *MIR*. The *MIR* terminator (508 bp) used to assess the terminator as described above was cloned into the *Sac*I and *Eco*RI sites of pBI121 containing *MIR* or *sMIR* after eliminating the *NOS* terminator fragment. In these constructs, *MIR* and *sMIR* gene expressions are driven by the constitutive *CaMV 35S* promoter.

The constructed plasmids were transferred into *Agrobacterium* as described above. Using *Agrobacterium* with the introduced plasmid, tomato (*Solanum lycopersicum* cv. Micro-Tom) plants were transformed as described by Sun et al. (2006b, 2007).

Genomic southern blot analysis

To confirm the copy number of the *MIR* or *sMIR* gene and the *neomycin phosphotransferase II* (*NPTII*) gene in transgenic tomato, genomic DNA was isolated using Maxwell® 16 DNA purification kits according to the manufacturer's

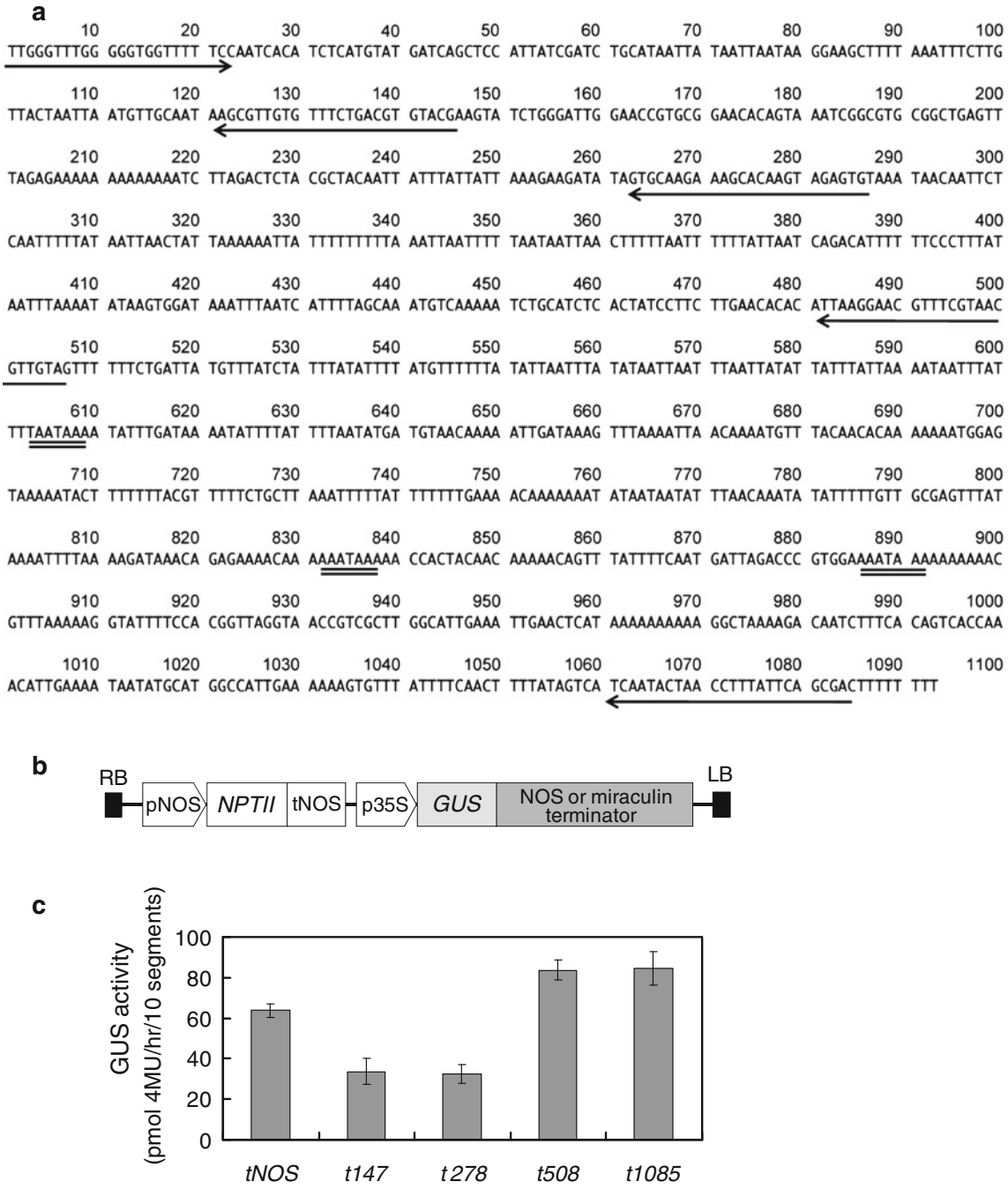

Fig. 1 Efficiency of gene expression with *MIR* terminators of various lengths in transfected tomato cotyledons. **a** Sequence alignment of the identified *MIR* terminator. *Arrows* indicate the positions and directions of the primers used for the transient assay. *Double lines* show the typical polyadenylation signal AATAAA. **b** T-DNA region of pBI121 used for the transient assay. The various lengths of the *MIR* terminator and *NOS* terminator were fused to the *GUS* gene under control of the *CaMV35S* promoter. LB and RB, the left and right borders of the T-DNA region, respectively; *pNOS, NOS* promoter; *NPTII, neomycin phosphotransferase* gene*; tNOS, NOS* terminator; *p35S CaMV 35S* promoter. **c** GUS activity in tomato cotyledons with *MIR* terminators of various lengths. GUS activity is expressed as the amount of the reaction product 4-methyl umbelliferone (4MU) per ten pieces of tomato cotyledon segments. *Vertical bars* show the standard error from three replications. *t147,* 147 bp terminator of the *MIR* terminator; *t278, t508* and *t1085* are abbreviated in the same manner as *t147*

protocol (Promega, Tokyo, Japan). Isolated genomic DNA (10 μg) was digested using the restriction enzyme *Xba*I (which cleaves only once outside the *MIR* gene), separated by electrophoresis on a 1% agarose gel and then transferred to a Hybond-N+ nylon membrane (GE Healthcare, Tokyo, Japan). Hybridization was performed overnight in high-SDS buffer (50% deionized formamide (v/v)), 5 × SSC, 7% SDS, 2% Blocking Regent (Roche, Tokyo, Japan), 50 mM

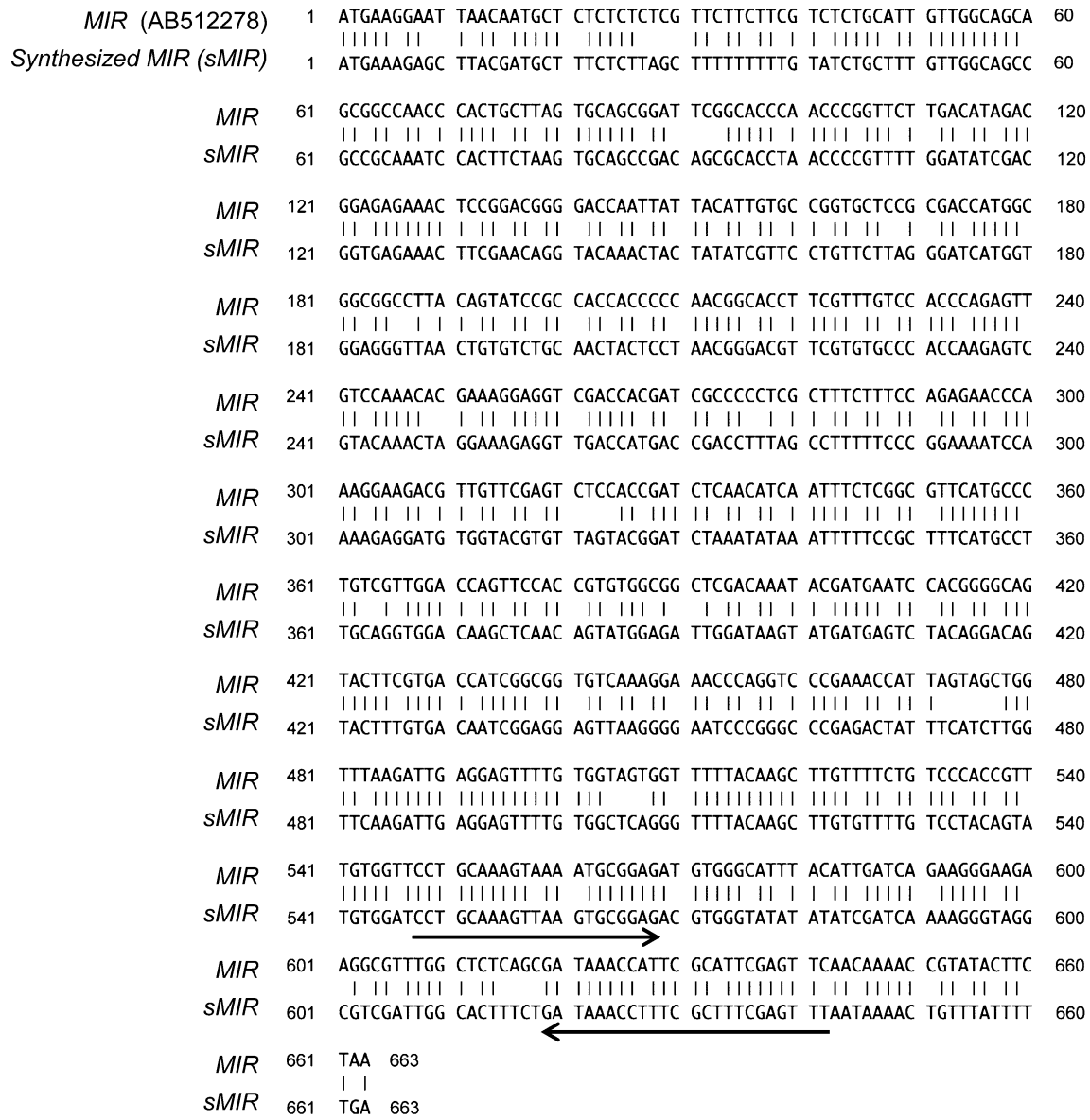

Fig. 2 DNA sequence comparison of the codon-modified *sMIR* coding region with the native *MIR* from miracle fruit. All codons of *sMIR* were optimized on the basis of the codon usage table for tomato. *Arrows* indicate the positions and directions of the primers used for qRT-PCR

sodium phosphate (pH 7.0), and 0.1% N-lauroylsarcosine sodium salt (w/v) containing a gene-specific DIG-labeled probe at 45°C. Probes were prepared with a PCR DIG Probe Synthesis Kit (Roche, Tokyo, Japan) following the manufacturer's protocol. The hybridization signal was detected by chemiluminescence using CDP-Star (Roche, Tokyo, Japan) followed by exposure in the LAS 4000 Mini Image Analyzer (Fujifilm Co. Ltd., Tokyo, Japan).

Immunoblot analysis and enzyme-linked immunosorbent assay (ELISA)

Tomato fruits were harvested during the red stage from *MIR* or *sMIR* and *NPTII* single-copy plants and ground to powder under liquid nitrogen. The protein was extracted as described previously by Hirai et al. (2010). Using the extracted protein, immunoblot analysis and ELISA were performed according to Sun et al. (2007) and Kim et al. (2010), respectively.

Quantitative reverse transcription-PCR (qRT-PCR) analysis

The expression levels of the *MIR* and *sMIR* transcripts in the transgenic tomato plants were determined by qRT-PCR. Total RNA was isolated from red fruits that were the same fruits used for immunoblot analysis and ELISA using the RNeasy Plant Mini Kit (Qiagen, Tokyo, Japan) with

RNase-free DNase (Qiagen, Tokyo, Japan). The first-strand cDNA was synthesized from extracted total RNA (0.5 μg) using the SuperScript™ III First-Strand Synthesis System (Invitrogen, Tokyo, Japan) with Oligo $(dT)_{20}$ primer. Ten-fold diluted first-strand cDNA was used as a template for the reaction with SYBR Premix Ex Taq II (Takara-Bio Inc., Otsu, Japan) on the Thermal Cycler Dice Real Time System TP800 (Takara-Bio Inc., Otsu, Japan) following the manufacturer's instructions. The qRT-PCR was subjected to 40 cycles of 95°C for 5 s and 57°C for 30 s. Relative quantification of *MIR* and *sMIR* gene expression was calculated by normalization to *Slubiquitin3* gene (GenBank accession number X58253) expression, which has been used as an internal control in tomato expression analysis (Hackel et al. 2006; Chincinska et al. 2008). Primer sequences were as follows: *MIR* forward, 5′-CCTGCAAA GTAAAATGCGGAGA-3′ and reverse, 5′-AACTCGAAT GCGAATGGTTTATC-3′; *sMIR* forward, 5′-CCTGCAAA GTTAAGTGCGGAGA-3′ and reverse, 5′-AACTCGAAA GCGAAAGGTTTATC-3′; and *ubiquitin* forward, 5′-CA CCAAGCCAAAGAAGATCA-3′ and reverse, 5′-TCAGC ATTAGGGCACTCCTT-3′. The *MIR* and *sMIR* primer sets were designed at the same sequence position although these primer sequences were not completely consistent.

Transcription termination of the *MIR* and *sMIR* transcripts harboring *NOS* or *MIR* terminators

To evaluate the influence of different terminators on transcription termination, RT-PCR was performed with several sets of primers specific to each gene using the GoTaq® Green Master Mix (Promega, Tokyo, Japan). cDNAs synthesized for qRT-PCR were used as a template. The amplification reactions consisted of 95°C for 3 min, 35 cycles of amplification (95°C for 30 s, 55°C for 30 s and 72°C for 2.5 min) and a final extension at 72°C for 7 min. The primers used are described in Table 1. *MIR* or *sMIR* mRNA's polyadenylation sites of each transgenic tomato from four different constructs and miracle fruit were detected using 3′-full RACE core set (TAKARA, Tokyo, Japan) as described in the manual. Total RNA was used to synthesize single-strand cDNA and amplified with PCR using gene-specific primers containing *Bam*HI or *Kpn*I sites for cloning: *Bam*HI-MIR, 5′-ACGGACGGA TCCAAAGGAAGACGTTGTTCGAGTCTC-3′; *Kpn*I-sMIR, 5′-ACGGACGGTACCTTTCATGCCTTGCAG GTGGACAAG-3′. Clones that confirmed the insertion by colony PCR were analyzed by the sequence with gene-specific primers: MIR-seq491F, 5′-AGATTGAGGAGT TTTGTGGTAGTGG-3′, sMIR-seq491F, 5′-AGATTGAG GAGTTTTGTGGCTCAGG-3′.

Results

Efficiency of GUS activity with *MIR* terminators of various lengths

We isolated a 1,953 bp fragment containing the *MIR* terminator, from which a sequence of 1,085 bp was used to evaluate its terminator function. The sequence included three typical polyadenylation signals (AATAAA) (Fig. 1a). Different length fragments of the terminator (146, 287, 508 and 1,085 bp) were bound downstream of the *GUS*-coding region and transferred to tomato cotyledon explants (Fig. 1b). When the *MIR* terminator fragment was over 508 bp, transient *GUS* expression (observed as a blue spot of GUS activity) was stronger than with the *NOS* terminator (data not shown). Similarly, GUS activity (as determined based on 4-methyl umbelliferone [4MU] accumulation) relative to that with the *NOS* terminator was higher with the 508 and 1,085 bp fragments and lower with the 146 and 287 bp fragments (Fig. 1c).

Production of transgenic tomato plants accumulating recombinant MIR protein

To assess the effects of codon optimization and use of the *MIR* terminator on MIR accumulation, four different constructs were created combining *MIR* or *sMIR* with the 508 bp *MIR* terminator sequence or the *NOS* terminator (*MIR-tNOS, MIR-tMIR, sMIR-tNOS* and *sMIR-tMIR*). In these constructs, *MIR* and *sMIR* gene expression are driven by the constitutive *CaMV 35S* promoter. Tomato cotyledons were transformed by infection with *Agrobacterium* containing the binary vector of each construct and then transformed tomato shoots were selected while rooting on medium including kanamycin. Kanamycin-resistant tomato lines were regenerated on selective medium, and the integration of the *MIR* gene in these plants was confirmed by southern blot analysis after selection of diploids using the ploidy test and by investigation of MIR accumulation via western blot analysis. Eventually, seven to 11 independent transgenic tomato lines were obtained as single-copy *MIR* or *sMIR* plus *NPTII* plants for each of the four constructs (Fig. 3).

Accumulation of recombinant MIR protein in transgenic tomato fruit

The concentration of recombinant MIR protein was measured with ELISA (Fig. 4). When the *MIR* terminator was used, the concentration in *MIR-tMIR* fruits was 1.5 times as high as that in the control (*MIR-tNOS*) fruits. Moreover, the highest effect on MIR concentration was detected in *sMIR-tMIR* fruits, in which the MIR concentration was 2.2 times

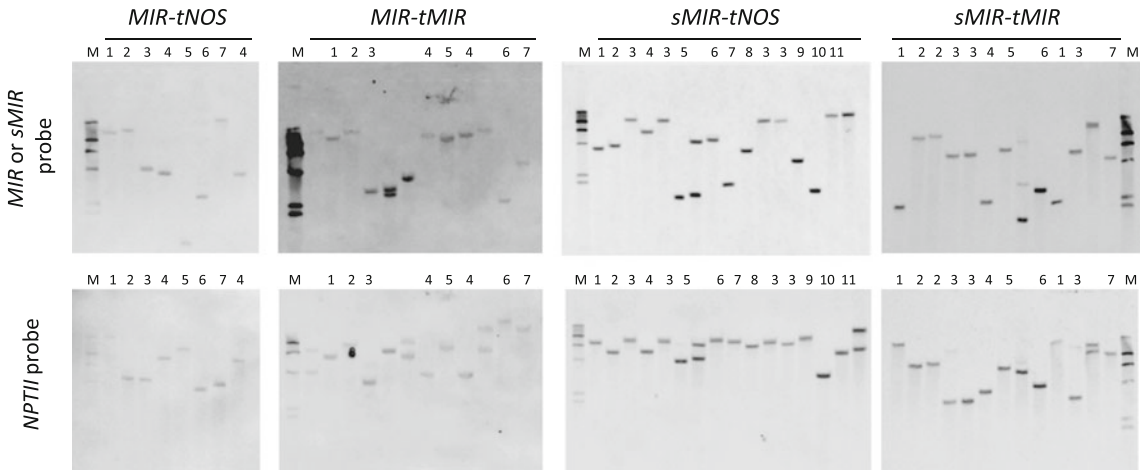

Fig. 3 Southern blot analysis of the *MIR*, *sMIR* and *NPTII* genes in transgenic tomatoes. DNA (10 μg) was digested with *Xba*I and detected with DIG-labeled probes from the coding sequences of *MIR*, *sMIR* or *NPTII*. *Lane numbers* show independent transgenic tomato lines (T_0) with each of the different constructs: *MIR-tNOS, MIR-tMIR, sMIR-tNOS, sMIR-tMIR*. *Tomato lines* with the same number are clones. *Upper* and *lower* photos show the data for *MIR* or *sMIR* with the *NPTII* marker, respectively. M, marker (DNA Molecular Weight Marker II, DIG-labeled; 125, 564, 2,027, 2,322, 4,361, 6,557, 9,416 and 23,130 bp)

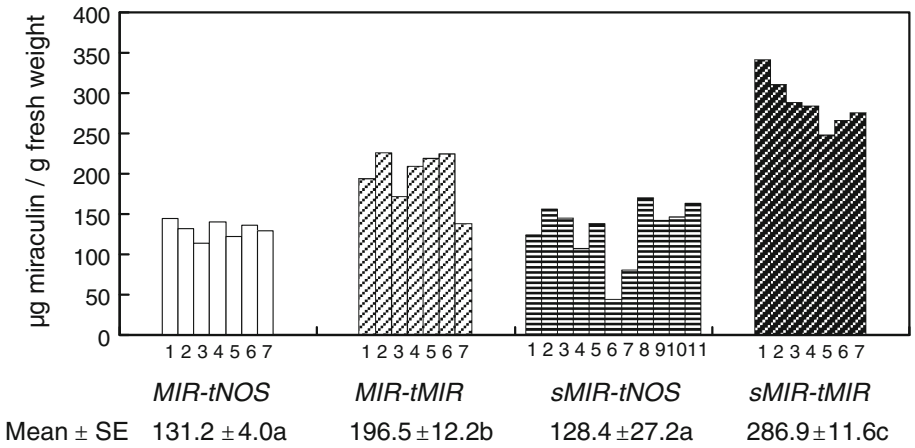

Fig. 4 MIR concentrations in transgenic tomatoes. Recombinant MIR protein was extracted from three to five red-ripe fruits from independent transgenic plants and the concentration was measured by ELISA. The *numbers* under the columns show the independent transgenic tomato lines identified as in Fig. 3. The presented data under the construct names are the mean ± standard errors (SE). Alphabets indicate significant differences based on Tukey–Kramer's Multiple Comparison test ($P < 0.01$)

higher than that in control fruit. In contrast, the MIR level in *sMIR-tNOS* fruits was almost the same as that in *MIR-tNOS* fruits. Western blot analysis of MIR protein levels reflected the ELISA results and also revealed that the molecular size of the signal from recombinant MIR was coincident with that of MIR purified from miracle fruit (data not shown).

Expression of the *MIR* and *sMIR* genes in transgenic tomato fruit

MIR and *sMIR* mRNA expression levels were detected by qRT-PCR. These sequences had 73% nucleotide homology (Fig. 2). Therefore, we were not able to design a primer set of completely as the same sequence for qRT-PCR. They had two mismatches in each of the forward and reverse primers. The two different primer sets were confirmed by amplification efficiency and characterization. When the *NPTII* primer set was used as a control, the standard curves of the *MIR* and *sMIR* primer sets exhibited almost the same slope and intercept as the standard curves of the *NPTII* primer set using either the *MIR* or *sMIR* plasmid as a template. Thus, these primer sets were used for qRT-PCR.

The expression levels of the mRNAs tended to be stronger when they were terminated by the *MIR* terminator than by the *NOS* terminator: compare *MIR-tMIR* to *MIR-tNOS* or *sMIR-tMIR* to *sMIR-tNOS* (Fig. 5). In addition, the expression of *MIR* mRNA was approximately two times

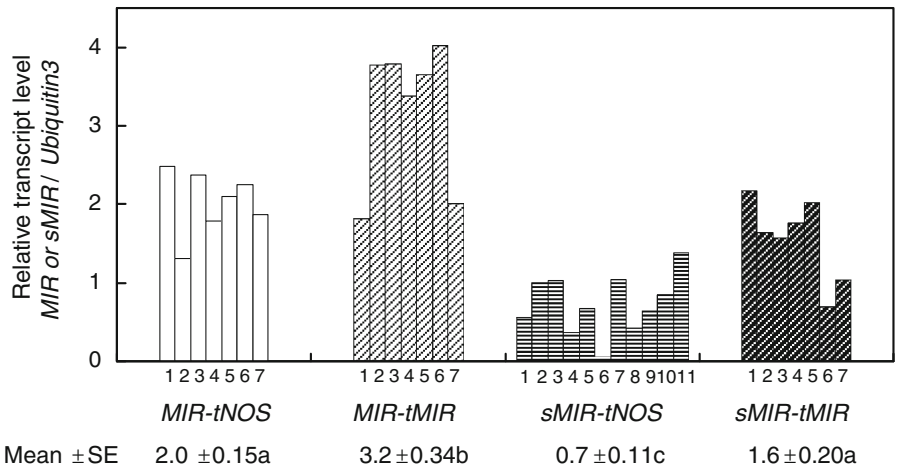

Fig. 5 Relative quantification of *MIR* and *sMIR* expression in transgenic tomatoes. *MIR* or *sMIR* mRNA was isolated from the same fruits used for Fig. 4, and the expression level was detected by real-time quantitative RT-PCR. The *numbers* under the columns show the independent transgenic tomato lines identified as in Fig. 3. The presented data under the construct names are the mean ± standard errors (SE). *Alphabets* indicate significant differences based on Tukey–Kramer's Multiple Comparison test ($P < 0.01$)

higher than that of *sMIR* mRNA when the terminator was the same.

Transcription termination with different terminators

To assess the effects of these terminators, read-through lengths were analyzed using RT-PCR. Transcription read-through was detected at the end of the *NOS* terminator with the *MIR-tNOS* and *sMIR-tNOS* mRNAs (Fig. 6a), and some of transcripts were detected around the left border. On the other hand, utilization of the *MIR* terminator did not induce read-through, at least up to the end of the *MIR* terminator. Subsequently, two sets of primers were designed inside the region of the *MIR* terminator and used for detailed investigation of the lengths of *MIR* and *sMIR* transcripts. mRNA read-through was verified at 146 bp on the *MIR* terminator with the *MIR-tMIR* and *sMIR-tMIR* constructs (Fig. 6b), and some read-through were also detected at 287 bp.

We also identified the polyadenylation sites using 3′ rapid amplification of cDNA ends (RACE). The major sites were at 57, 151 and 166 bp in *MIR-tNOS* mRNA and, at 99 and 111 bp in *MIR-tMIR* mRNA with referrence to bases downstream from the stop codon (Fig. 7). In *sMIR-tNOS* and *sMIR-tMIR* mRNAs, the major sites were almost at the same position in *MIR-tNOS* and *MIR-tMIR*, respectively. The sites of native *MIR* mRNA from miracle fruit were at 104 and 105 bp close to those of *MIR-tMIR* and *sMIR-tMIR*.

Discussion

The terminator sequence plays a key role in transcript termination, mRNA stability and mRNA modifications such as capping, splicing and polyadenylation (Proudfoot 2004; Gilmartin 2005; Desai et al. 2010). Consequently, it also regulates by these functions the level of mRNA expression and protein accumulation. In particular, polyadenylation is important for the regulation of mRNA stability, transportation and translation (Jackson and Standart 1990; Zarudnaya et al. 2003). We identified and analyzed a 1,086 bp fragment containing the *MIR* terminator. Three AAUAAA motifs—this motif being a well known, typical polyadenylation signal—were found in this sequence. GUS activity in a transient assay was almost equivalent with *t507* and *t1086* even though the AAUAAA signal is not contained in the *t507* sequence (Fig. 1). Plants can recognize AAUAAA-like signals with a high degree of sequence variation (Rothnie et al. 1994), and more than ten AAUAAA-like signal sequences were confirmed in the *t507* sequence. Some of these signals may contribute to polyadenylation.

Different types of terminators influence gene expression levels (Ingelbrecht et al. 1989; Nagaya et al. 2010). A terminator derived from the *heat shock protein 18.2* gene increased mRNA expression levels 2.5-fold and 1.5-fold over that with the *NOS* terminator in transient assays using protoplasts of *Arabidopsis* and rice, respectively (Nagaya et al. 2010). In this study, the expression levels of both the *MIR* and the *sMIR* genes were enhanced using the *MIR* terminator instead of the *NOS* terminator (Fig. 5). This result was also reflected in recombinant MIR protein levels (Fig. 4). We used RT-PCR to confirm that mRNAs of varying lengths were detected with the *NOS* terminator. However, this variation was considerably less using the *MIR* terminator. Additionally investigation of the polyadenylation sites indicated that mRNAs analyzed for the

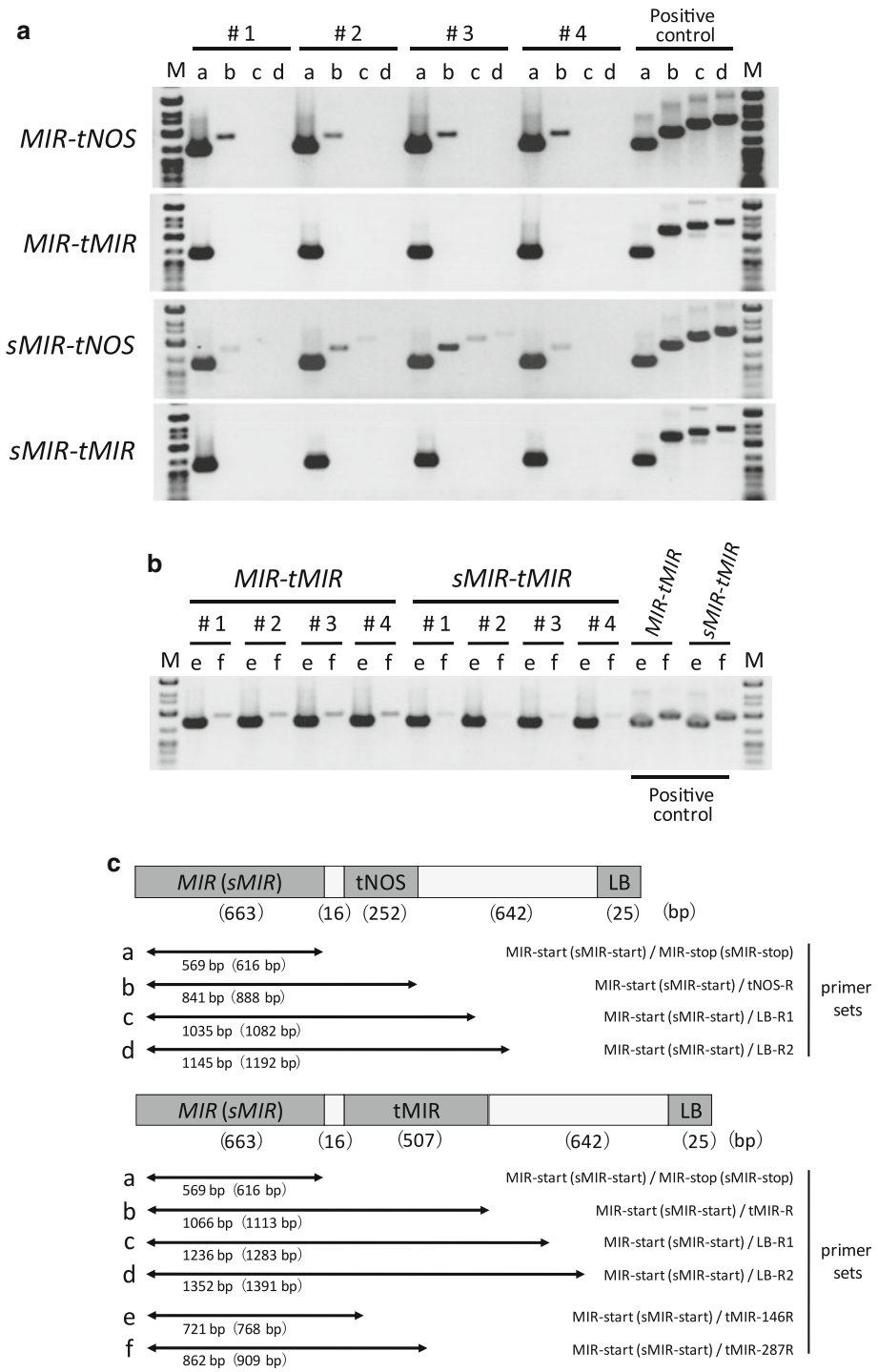

Fig. 6 Transcription termination of the *MIR* and *sMIR* transgenes. The cDNA samples used in Fig. 5 were analyzed by PCR using sets of primers (Table 1). The *numbers* (#1 to #4) show the independent transgenic tomato lines identified as in Fig. 3. Plasmid DNA was used as a positive control template. M, marker (Gene Ladder Wide 1, Nippon gene). **a** Detection of read-through transcripts with four different sets of primers. Small letters (*a*, *b*, *c*, *d*) above the photos show the primer sets described in Fig. 6c. **b** Detection of read-through transcripts from the *MIR* terminator. *Small letters* (*e*, *f*) above the photos show the primer sets described in Fig. 6c. **c** Primer design and size of PCR products. *Numbers* under the constructs show the nucleotide size of each region. The *small letters* on the left side (*a*–*f*) depict the sets of primers used and correspond to Fig. 6a, b. The size of the product from the *MIR* gene is indicated under each *arrow* and the size of the product from the *sMIR* gene is in *parentheses*. *tNOS, NOS* terminator; *tMIR, MIR* terminator; LB, left border

3′-UTR from *MIR* terminator were shorter than the major one (166 bp) from *NOS* terminator and the size variation was less like that from miracle fruit than that of *NOS* terminator. On the other hand, we were not able to find the longer read-through mRNAs such around 268 bp at *NOS* terminator (Fig. 6a, lane b) and 146 bp at *MIR* terminator (Fig. 6b, lane e) with reference to bases downstream from the stop codon among the detected polyadenylation sites. The longer mRNAs identified with RT-PCR (Fig. 6a, b) might be minor. These results indicate that the capacity for transcription termination of the *MIR* terminator is greater than that of the *NOS* terminator. The increased MIR accumulation in this study may be the result of a change in mRNA stability due to either transcript termination or the *MIR* terminator sequence itself that further influences the production of protein.

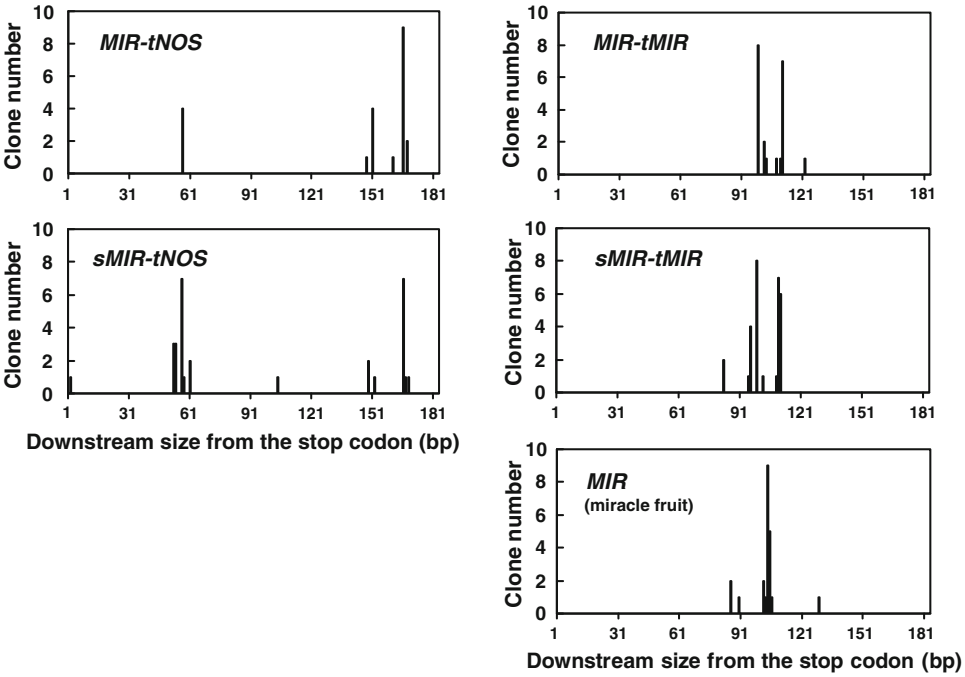

Fig. 7 Polyadenylation sites of transcripts from *MIR* and *sMIR* transgenes and native *MIR* gene of miracle fruit. 3′-RACE was performed using cDNA prepared from #1 and #2 transgenic lines for each construct. The resulting PCR products were cloned and sequenced. Polyadenylation sites were counted downstream from the transcriptional stop codon. The number of investigated clones was as follows: *MIR-tNOS*, 21 clones; *MIR-tMIR*, 21 clones; *sMIR-tNOS*, 30 clones, *sMIR-tMIR*, 30 clones, native *MIR* from miracle fruit, 22 clones

The *NOS* terminator from the Ti plasmid of *Agrobacterium* is universally used in various expression vectors when genes are transformed into a plant. In fact, this terminator is used in some commercialized transgenic crops given herbicide-tolerance and/or resistance to a pest, including soybean, corn, rape, cotton and potato. It is well known that failure to terminate transcription, and read-through of transcription occurs beyond the *NOS* terminator (Windels et al. 2001). Read-through transcripts may contain an open reading frame (ORF) other than the target gene and sometimes unknown peptides from the ORF are produced (Rang et al. 2005). In the case of transgenic plants for human consumption, the existence of an additional peptide outside of the target protein is not suitable for commercial use. Read-through can also cause transcriptional interference of genes located downstream of the terminator (Ingelbrecht et al. 1991). Thus, the function of transcript termination is crucial not only for tuning expression levels but also for avoiding improper peptide production and transcriptional interference. In addition to these points, the *MIR* terminator is derived from an edible plant, the miracle fruit. Our results have provided one of the most useful terminators for commercial use.

Optimization of codon usage is frequently used when a gene introduced into a host plant is derived from another organism such as fungi (Xue et al. 2003; Peng et al. 2006), bacteria (Perlak et al. 1991) or animals (Rouwendal et al. 1997) because codon preferences are quite different in these kingdoms. In plants, it is also different not only between monocots and dicots but also between the nucleus and the plastid of the same plant (Batard et al. 2000; Kawabe and Miyashita 2003; Liu and Xue 2005). Therefore, modifying codons to suit the host may significantly improve the production of a target protein, especially when rare codons are used in the heterologous gene. We optimized the codons of the *MIR* gene to make it suitable for tomato. As a result, the production of recombinant MIR protein in *sMIR-tMIR* fruits was higher than that in *MIR-tMIR* fruits, although the expression of *sMIR* mRNA was low compared to that of *MIR-tMIR*, suggesting that the improvement in translation efficiency exceeded the decline in transcription efficiency (Fig. 4). Similar results were observed between *sMIR-tNOS* and *MIR-tNOS* fruits. The expression of *MIR* mRNA in *MIR-tNOS* was higher than that of *sMIR* mRNA in *sMIR-tNOS*, but the MIR productivity was almost the same. From another standpoint, the *sMIR* transcript levels were approximately half than those of *MIR* using either the *NOS* or the *MIR* terminator; however, the impact on translation of codon optimization was higher when using the *MIR* terminator than using the *NOS* terminator (Figs. 4, 5). These results imply that mRNA terminated by the *NOS* terminator causes a decline in translational efficiency.

In this study, the *sMIR* transcript levels of *sMIR-tNOS* and *sMIR-tMIR* fruits reduced compared to that of *MIR*. Some reports indicate that a sequence of 5′-UTR and the amino acid sequence of the first exon coding region influence transcription initiation, transcriptional efficiency and mRNA stability (Chiba et al. 1999; Gutiérrez et al. 1999; Suzuki et al. 2001; Matsuura et al. 2008). However, it is unknown if ORF sequences themselves have any effect on transcription. The unknown character of ORF sequences

may be responsible for the decline in the amount of *sMIR* transcripts. Another possibility is that the secondary structure of mRNA may influence mRNA stability. The codons of the *MIR* gene were thoroughly modified. Therefore, there is a possibility that some changed sequences were critical for transcriptional efficiency and mRNA stability. Either way, further study is required to understand this mechanism. If the transcript level can be increased by further optimization of the *MIR* gene, higher production of recombinant MIR protein might be achieved using the *MIR* terminator.

In conclusion, we succeeded at producing transgenic tomatoes with recombinant MIR at concentrations up to 340 μg/g fresh weight using the *MIR* terminator and codon optimization of the *MIR* gene, although the concentration was almost 100 μg/g fresh weight in our previous study (Sun et al. 2007; Hirai et al. 2010; Yano et al. 2010). Additionally, we suggest that the identified *MIR* terminator is useful for increasing the level of transcription, improving translational efficiency by codon modification and improving the quality of the mRNA. These factors play a key role in the final productivity.

Acknowledgments We thank members of the Ezura laboratory for helpful discussions. We thank Sun HJ, Uchi S and Yano M for cloning of the miraculin terminator region. This study was supported by the project "Development of Fundamental Technologies for the Production of High-value Materials Using Transgenic Plants," by the Ministry of Economy, Trade, and Industry of Japan to H.E. Micro-Tom seeds (TOMJPF00001) were provided by the Gene Research Center, University of Tsukuba, through the National BioResource Project (NBRP) of the MEXT, Japan.

References

Batard Y, Hehn A, Nedelkina S, Schalk M, Pallett K, Schaller H, Werck-Reichhart D (2000) Increasing expression of P450 and P450-reductase proteins from monocots in heterologous systems. Arch Biochem Biophys 379:161–169. doi:10.1006/abbi.2000.1867

Chiba Y, Ishikawa M, Kijima F, Tyson RH, Kim J, Yamamoto A, Nambara E, Leustek T, Wallsgrove RM, Naito S (1999) Evidence for autoregulation of cystathionine γ-synthase mRNA stability in *Arabidopsis*. Science 286:1371–1374. doi:10.1126/science.286.5443.1371

Chincinska IA, Liesche J, Krugel U, Michalska J, Geigenberger P, Grimm B, Khun C (2008) Sucrose transporter StSUT4 from potato affects flowering, tuberization, and shade avoidance response. Plant Physiol 146:515–528. doi:10.1104/pp.107.112334

Daniell H, Streatfield SJ, Wycoff K (2001) Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. Trends Plant Sci 6:219–226. doi:10.1016/S1360-1385(01)01922-7

Deblaere R, Bytebier B, De Greve H, Deboeck F, Schell J, van Montagu M, Leemans J (1985) Efficient octopine Ti plasmid derived vectors for *Agrobacterium*-mediated gene transfer in plants. Nucleic Acids Res 13:4777–4785

Desai PN, Shrivastava N, Padh H (2010) Production of heterologous proteins in plants: Strategies for optimal expression. Biotechnol Adv 28:427–435. doi:10.1016/j.biotechadv.2010.01.005

Giddings G, Allison G, Brooks D, Carter A (2000) Transgenic plants as factories for biopharmaceuticals. Nat Biotechnol 18:1151–1155. doi:10.1038/81132

Gilmartin GM (2005) Eukaryotic mRNA 3′ processing: a common means to different ends. Genes Dev 19:2517–2521. doi:10.1101/gad.1378105

Gustafsson C, Govindarajan S, Minshull J (2004) Codon bias and heterologous protein expression. Trends Biotechnol 22:346–353. doi:10.1016/j.tibtech.2004.04.006

Gutiérrez RA, Maclntosh GC, Green PJ (1999) Current perspectives on mRNA stability in plants: multiple levels and mechanisms of control. Trend Plant Sci 4:429–438. doi:10.1016/S1360-1385(99)01484-3

Hackel A, Schauer N, Carrari F, Fernie AR, Grimm B, Kuhn C (2006) Sucrose transporter LeSUT1 and LeSUT2 inhibition affects tomato fruit development in different ways. Plant J 45:180–192. doi:10.1111/j.1365-313X.2005.02572.x

Hirai T, Fukukawa G, Kakuta H, Fukuda N, Ezura H (2010) Production of recombinant miraculin using transgenic tomatoes in a closed cultivation system. J Agric Food Chem 58:6096–6101. doi:10.1021/jf100414v

Ingelbrecht ILW, Herman LMF, Dekeyser RA, Van Montagu MC, Depicker AG (1989) Different 3′ end regions strongly influence the level of gene expression in plant cells. Plant Cell 1:671–680. doi:10.1105/tpc.1.7.671

Ingelbrecht ILW, Breyne P, Vancompernolle K, Jacobs A, Van Montagu MC, Depicker AG (1991) Transcriptional interference in transgenic plants. Gene 109:239–243. doi:10.1016/0378-1119(91)90614-H

Jackson RJ, Standart N (1990) Do the poly (A) tail and 3′ untranslated region control mRNA translation? Cell 62:15–24. doi:10.1016/0092-8674(90)90235-7

Jefferson RK, Kavanagh TA, Bevan MW (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J 6:3901–3907

Kawabe A, Miyashita NT (2003) Patterns of codon usage bias in three dicot and four monocot plant species. Genes Genet Syst 78:343–352. doi:10.1266/ggs.78.343

Kim YW, Kato K, Hirai T, Hiwasa-Tanase K, Ezura H (2010) Spatial and development profiling of miraculin accumulation in transgenic tomato fruits expressing the miraculin gene constitutively. J Agric Food Chem 58:282–286. doi:10.1021/jf9030663

Liu Q, Xue Q (2005) Comparative studies on codon usage pattern of chloroplasts and their host nuclear genes in four plant species. J Genet 84:55–62

Mason HS, Warzecha H, Mor T, Arntzen CJ (2002) Edible plant vaccines: application for prophylactic and molecular medicine. Trends Mol Med 8:324–329. doi:10.1016/S1471-4914(02)02360-2

Matsuura H, Shinmyo A, Kato K (2008) Preferential translation mediated by Hsp81–3 5′-UTR during heat shock involves ribosome entry at the 5′-end rather than an internal site in *Arabidopsis* suspention cells. J Biosci Bioeng 105:39–47

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue culture. Physiol Plant 15:473–497. doi:10.1111/j.1399-3054.1962.tb08052.x

Murray EE, Lotzer J, Eberle M (1989) Codon usage in plant genes. Nucleic Acids Res 17:477–498

Nagaya S, Kawamura K, Shinmyo A, Kato K (2010) The *HSP* terminator of *Arabidopsis thaliana* increases gene expression in plant cells. Plant Cell Physiol 51:328–332. doi:10.1093/pcp/pcp188

Peng RH, Yao QH, Xiong AS, Cheng ZM, Li Y (2006) Codon-modifications and an endoplasmic reticulum-targeting sequence additively enhance expression of an *Aspergillus* phytase gene in transgenic canola. Plant Cell Rep 25:124–132. doi:10.1007/s00299-005-0036-y

Perlak FJ, Fuchs RL, Dean DA, McPherson SL, Fischhoff DA (1991) Modification of the coding sequence enhances plant expression of insect control protein genes. Proc Natl Acad Sci USA 88:3324–3328

Proudfoot N (2004) New perspectives on connecting messenger RNA 3′ end formation to transcription. Curr Opin Cell Biol 16:272–278. doi:10.1016/j.ceb.2004.03.007

Rang A, Linke B, Jansen B (2005) Detection of RNA variants transcribed from the transgene in Roundup Ready soybean. Eur Food Res Technol 220:438–443. doi:10.1007/s00217-004-1064-5

Rogers SO, Bendich AJ (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Plant Mol Biol 5:69–76. doi:10.1007/BF00020088

Rothnie HM, Reid J, Hohn T (1994) The contribution of AAUAAA and the upstream element UUUGUA to the efficiency of mRNA 3′-end formation in plants. EMBO J 13:2200–2210

Rouwendal GJ, Mendes O, Wolbert EJ, de Boer AD (1997) Enhanced expression in tobacco of the gene encoding green fluorescent protein by modification of its codon usage. Plant Mol Biol 33:989–999. doi:10.1023/A:1005740823703

Shen WJ, Forde BG (1989) Efficient transformation of *Agrobacterium* spp. by high voltage electroporation. Nucleic Acids Res 17:8385

Sugaya T, Yano M, Sun HJ, Hirai T, Ezura H (2008) Transgenic strawberry expressing a taste-modifying protein, miraculin. Plant Biotechnol 25:329–333

Sun HJ, Cui ML, Ma B, Ezura H (2006a) Functional expression of the taste-modifying protein, miraculin, in transgenic lettuce. FEBS Lett 580:620–626. doi:10.1016/j.febslet.2005.12.080

Sun HJ, Uchii S, Watanabe S, Ezura H (2006b) A highly efficient transformation protocol for Micro-Tom, a model cultivar for tomato functional genomics. Plant Cell Physiol 47:426–431. doi:10.1093/pcp/pci251

Sun HJ, Kataoka H, Yano M, Ezura H (2007) Genetically stable expression of functional miraculin, a new type of alternative sweetener, in transgenic tomato plants. Plant Biotechnol J 5:768–777. doi:10.1111/j.1467-7652.2007.00283.x

Suzuki A, Shirata Y, Ishida H, Chiba Y, Onouchi H, Naito S (2001) The first exon coding region of cystathionine γ-synthase gene is necessary and sufficient for downregulation of its own mRNA Accumulation in transgenic *Arabidopsis thaliana*. Plant Cell Physiol 42:1174–1180

Theerasilp S, Kurihara Y (1988) Complete purification and characterization of the taste-modifying protein, miraculin, from miracle fruit. J Biol Chem 263:11536–11539

Twyman RM, Stoger E, Schillberg S, Christou P, Fischer R (2003) Molecular farming in plants: host systems and expression technology. Trends Biotechnol 21:570–578. doi:10.1016/j.tibtech.2003.10.002

Windels P, Taverniers I, Depicker A, Van Bockstaele E, De Loose M (2001) Characterisation of the Roundup Ready soybean insert. Eur Food Res Technol 213:107–112

Xue GP, Patel M, Johnson JS, Smyth DJ, Vickers CE (2003) Selectable marker-free transgenic barley producing a high level of cellulose (1, 4-β-glucanase) in developing grains. Plant Cell Rep 21:1088–1094. doi:10.1007/s00299-003-0627-4

Yano M, Hirai T, Kato K, Hiwasa-Tanase K, Fukuda N, Ezura H (2010) Tomato is a suitable material for producing recombinant miraculin protein in genetically stable manner. Plant Sci 178:469–473. doi:10.1016/j.plantsci.2010.02.016

Zarudnaya MI, Kolomiets M, Potyahaylo AL, Hovorun DM (2003) Downstream elements of mammalian pre-mRNA polyadenylation signals: primary, secondary and higher-order structures. Nucleic Acids Res 31:1375–1386