GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

A novel synthetic *Cry1Ab* gene resists rice insect pests

F.S. Song[1,2], D.H. Ni[2], H. Li[2,3], Y.B. Duan[2], Y.C. Yang[2], J.L. Ni[2], X.Z. Lu[2], P.C. Wei[2,3], L. Li[2] and J.B. Yang[2]

[1]College of Life Sciences, Anhui Agricultural University, Hefei, China
[2]Key Laboratory of Rice Genetics Breeding of Anhui Province, Rice Research Institute, Anhui Academy of Agricultural Sciences, Hefei, China
[3]Institute of Agricultural Engineering, Anhui Academy of Agricultural Sciences, Hefei, China

Corresponding author: J.B. Yang
E-mail: yjianbo@263.net

Genet. Mol. Res. 13 (2): 2394-2408 (2014)
Received August 13, 2013
Accepted December 10, 2013
Published April 3, 2014
DOI http://dx.doi.org/10.4238/2014.April.3.12

ABSTRACT. A few insect control genes of *Bacillus thuringiensis* have been modified successfully to increase the expression in plants by replacing rare codons, increasing GC content, and avoiding the DNA elements that could cause premature transcription termination, mRNA instability, and potential methylation. However, the modification process was intricate and often confused researchers. In this study, we adopted a simple method to modify *Cry1Ab* only by individually replacing its amino acid sequence with corresponding rice-preferred codons based on analysis of 92,188 coding DNA sequences. Unexpectedly, all elements of A+T richness, which terminate or destabilize transcription in plants, were avoided in the newly designed *mCry1Ab*. However, *mCry1Ab* had 2 notable features: less synonymous codons and high GC content. *mCry1Ab* only employed 22 of the 61 codons to encode protein and had an enhanced GC content of 65%. The increase in GC content caused abundant potential methylation signals to emerge in *mCry1Ab*. To test whether *mCry1Ab* could be expressed in rice, we transferred it into *Oryza japonica* variety

 ©FUNPEC-RP www.funpecrp.com.br

Wanjing97. Insect bioassays revealed that transgenic plants harboring this gene driven by 2 promoters, *CaMV35S* and *OsTSP I*, were highly resistant to rice leaffolder (*Cnaphalocrocis medinalis*). Analysis of R_0 to R_2 generation plants indicated that the *mCry1Ab* was inherited stably by the progeny. Our study provided a simple modified method for expressing exogenous genes in rice and confirmed that less synonymous codons and high GC content do not affect transgene expression in rice.

Key words: Gene modification; GC content; Synonymous codons; *Cry1Ab*; Insect resistance

INTRODUCTION

The toxins produced by *Bacillus thuringiensis* (*Bt*) kill some major insect pests but are safe for vertebrates and other organisms (Siegel, 2001; Park et al., 2012). *Bt* toxins have been used as sprayable bioinsecticides in agricultural production for decades (García-Robles et al., 2013). However, due to limited field stability and an inability to reach cryptic insects, *Bt* insecticides still hold only a small share of the insecticide market (Ferré and Van Rie, 2002). With the development of transgenic techniques, *Bt* genes can be introduced into plant genomes for *in vivo* expression to solve the problems of instability and inability to reach to cryptic insects. Nevertheless, it was proven that directly transferring wild-type *Bt* genes into plants resulted in poor expression, with the toxin comprising less than 0.005% of the total proteins in the plant, making transgenic plants still susceptible to insects even if the transgenes were driven by a strong plant promoter (Barton et al., 1987; Vaeck et al., 1987; Perlak et al., 1991).

To improve the *Bt* gene expression level, Perlak et al. (1991) modified *Cry1Ab* and *Cry1Ac* by using plant-biased codons to eliminate sequences such as potential polyadenylation signal sequences, ATTTA sequences, and A+T-rich and G+C-rich regions (N > 4 bases), increasing their expression levels 10-100 fold higher than those of the wild-type genes. Fujimoto et al. (1993) optimized another truncated *Cry1Ab* by replacing A and T of the third position of triplet codons with a randomly chosen G or C to make the relative frequency of codon groups (XXG, XXC, XXA, and XXT, in which X is any base) in the modified gene similar to the 12 rice genes that are highly expressed in leaves and transgenic rice that is resistant to striped stem borer and leaffolder. Subsequently, based on a similar design, several other *Bt* genes such as hybrid *Cry1Ab*/*Cry1Ac*, *Cry2A**, *Cry1C**, and *cry1Ca1* were successfully synthesized in rice (Tu et al., 1998; Chen et al., 2005; Tang et al., 2006; Zaidi et al., 2009). However, the modification process was intricate and often confused researchers.

In this study, we adopted a simple method to modify *Cry1Ab* only by individually replacing its amino acid sequence with corresponding rice-preferred codons, and the newly designed gene was called *mCry1Ab*. Unexpectedly, *mCry1Ab* easily eliminated all elements of A+T richness that were common in wild-type *Cry1Ab* - e.g., AATAAA, AATAATT, and ATTTA-which terminate or destabilize transcription in plants. Compared with other modified *Bt* gene versions, *mCry1Ab* uses only 1 codon for most amino acids. Moreover, the GC content of *mCry1Ab* was enhanced to 65%, reaching the maximum value. The increase in GC content causes sequences of continuous G+C (> 6 bases) to emerge frequently in *mCry1Ab*, as many as 23 times compared to 2 times in the original *Cry1Ab*. It was proposed that sequences of continuous G+C should be avoided in plants and other eukaryotes because they may be regulated

by DNA methylation, which can result in gene silencing (Murray et al., 1989; Bender, 2004). Additionally, Perlak et al. (1991) avoided G+C (> 4 bases) sequences in their modified *Bt* genes. To test whether *mCry1Ab* could be expressed in rice, we transferred it into *Oryza japonica* variety Wanjing97 and detected its expression level and ability to confer insect resistance.

MATERIAL AND METHODS

Analysis of codon usage in *Bt* and rice

The codon information of rice and *Bt* was acquired from the codon usage database (http://www.kazusa.or.jp/codon) (Nakamura et al., 2000). Rice had 92,188 coding DNA sequences (CDSs) with a total of 34,132,283 codons, and *Bt* had 496 CDSs with a total of 211,286 codons. The frequency of synonymous codons for each amino acid was calculated, and the codons with the highest usage frequency among synonymous codons were chosen as preferred codons.

Design and synthesis of *mCry1Ab*

Cry1Ab, which is 1845 bp in length, was confirmed to be expressed efficiently in its modified version in dicots by Perlak et al. (1991). In this study, amino acids in the Cry1Ab protein were replaced individually with corresponding rice-preferred codons to form a new *Cry1Ab* gene called *mCry1Ab*, which was synthesized by Invitrogen Life Technologies Co., Ltd. (Shanghai, China) and transferred into the pMD18 vector to form the pMD18-*mCry1Ab* vector.

Construction of the expression vector and transformation

The pMD18-*mCry1Ab* vector (*Escherichia coli* JM110 strain) was extracted using the Axygen plasmid extraction kit and was used as the polymerase chain reaction (PCR) template. *mCry1Ab* was amplified using the primers *mCry1Ab*-F_1 (5'-GATAACCTGAAGACTCCATGGACAACAACCCGAACATCAACG-3') and *mCry1Ab*-R_1 (5'-CCAGCATTGAAGACTCGTCACCTCAGTACTCGGCCTCGAAGGTCACC-3'). The amplification reagents were as follows: 5 µL 10X PCR buffer, 4 µL 25 mM $MgCl_2$, 4 µL 2.0 mM dNTPs, 2 µL 10 µM primers, 0.4 µL 5 U/µL *Taq* enzyme, and 4 µL 25 ng/µL DNA template; finally, the total volume was brought up to 50 µL with ddH_2O. The PCR was performed at 94°C for 5 min; 30 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 2 min; and a final step of 72°C for 7 min. The PCR products were digested with *Bbs*I, purified, and used to replace *Gus* of pCAMBIA1305.1, forming the expression vector p*35s*::*mCry1Ab* (Figure 1a).

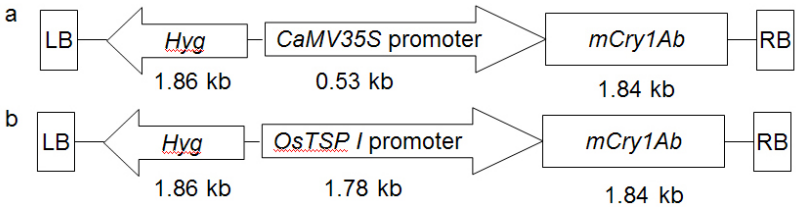

Figure 1. Transfer DNA regions of the transformed expression vector. **a.** *mCry1Ab* was controlled by the *CaMV35S* promoter and nos terminator. **b.** *mCry1Ab* gene was controlled by the *OsTSPI* promoter and the nos terminator. Hygromycin (Hyg) was used as a selective gene under the control of the *CaMV35S* promoter and the *CaMV35S* polyA tail.

OsTSP I is a rice promoter that was cloned by the Rice Research Institute, Anhui Academy of Agricultural Sciences, China, and was fused into the p*OsTSP I::GUS* vector. The *mCry1Ab* fragment was amplified using the primers *mCry1Ab*-F_2 (5'-GATAACCTGAAGAC TCCATGGACAACAACCCGAACATCAACG-3') and *mCry1Ab*-R_2 (5'-CTTAGCGCTCTC AGTACTCGGCCTCGAAGGTCACC-3'). The p*OsTSP I::Gus* vector was double digested with *Nco*I and *Afe*I to remove *Gus*. The PCR amplification fragment that was digested with *Bbs*I and *Afe*I was ligated with the fragment of the p*OsTSP I::Gus* vector without *Gus* to form the p*OsTSPI::mCry1Ab* vector (Figure 1b).

EHA105 strains of *Agrobacterium tumefaciens* harboring the p*35s*::*mCry1Ab*, p*OsTSPI::mCry1Ab*, and pCAMBIA1305.1 vectors were used for the rice transformation experiments. The elite japonica rice Wanjing97 was used as the recipient. Transformation procedures were performed as described by Duan et al. (2012) with the modification of using 50 mg/L hygromycin as the selectable agent instead of mannose.

PCR analysis

mCry1Ab-specific primers (forward 5'-TGTCCCTCTTCCCGAACTACG-3' and reverse 5'-GAGGTGCGGGCTCCTGATGG-3') were designed for PCR amplification, producing a target product of 158 bp. Total genomic DNA was extracted according to Li et al. (2013). The PCR was performed in a total volume of 25 µL containing 2 µL 25 ng/µL genomic DNA, 1 µL 10 µM primers, 1 µL 2.0 mM dNTPs mix, 2.5 µL 10X PCR buffer (containing 500 mM KCl, 100 mM Tris-HCl, pH 8.3, and 30 mM $MgCl_2$), and 1 U *Taq* polymerase. The amplification program was performed with an initial denaturation step at 94°C for 5 min; 35 cycles of denaturation at 94°C for 30 s, annealing at 58°C for 30 s, and extension at 72°C for 30 s; and a final extension at 72°C for 7 min. The DNA fragments were separated on 2% agarose gels and stained with ethidium bromide.

Cry1Ab/Ac strips and enzyme-linked immunosorbent assay (ELISA)

The Cry1Ab/Ac test strip (Beijing Silver Soil Biotechnology Co., Ltd, China) can sensitively detect 2 proteins, Cry1Ab and Cry1Ac, in plants. A rice leaf (5 cm in length) was put into a 1.5-mL centrifuge tube with 50 µL purified water, and it was crushed with a tip until the liquid was green or turbid. The test strip was put into the samples to ensure that the liquid level was below the Max line of the test strip. After incubating the samples for 1 min, the test strip was removed and placed for 1 min, and then the detection line was observed.

The expression of Cry1Ab toxic protein in transgenic plants was quantitatively analyzed using the Cry1Ab/Cry1Ac plate kit purchased from Envirologix Corporation (USA) according to manufacturer instructions. Leaves of transgenic plants at the tillering stage were ground into a powder using liquid nitrogen, and 0.02 g powder was weighed into 1.5-mL centrifuge tubes with 500 µL extraction buffer. Samples were diluted 400-fold after serious oscillation, and 50 µL each sample was used for detection. The plants that were transformed with the blank vector pCAMBIA1305.1 were used as the negative control. The standard curve was constructed using a concentration gradient (0.02, 0.016, 0.012, 0.008, 0.004, and 0 ng) for quantitative analysis of samples.

Bioassay in the laboratory

To test whether Cry1Ab proteins produced in the transgenic plants had insecticidal activity toward the rice leaffolder, an insect bioassay was conducted using the test tube method according to Qiu et al. (2010). Rice leaves at the tillering stage were cut into 10-cm-long sections. Five leaf sections were placed into a test tube filled with 1% agar to a depth of 2 cm. Leaffolder larvae at the second to third instar stages were collected from the rice field, and 5 larvae were released into each test tube. Test tubes were sealed with plastic film to prevent the larvae from escaping. Small holes were made in the plastic film by poking it with a needle. The feeding assay was conducted in a growth chamber at a temperature of 28°C, a relative humidity of 85%, and a photoperiod of 12 h. Damage to the leaf tissues and larval mortality were observed and photographed 4 days after infestation. The plants that were transformed with the blank vector pCAMBIA1305.1 were used as the negative control.

Evaluation of insect resistance in the field

The transgenic plants were planted in a net house at the Rice Research Institute, Anhui Academy of Agricultural Science, China, and surrounded with wild-type Wanjing97 as a negative control. No pesticide was applied during the growth period. The insect occupancy of negative and wild-type rice was observed every 7 days. When every negative control plant showed insect damage, the insect occurrence of transgenic plants was investigated. For R_0 and R_1 transgenic plants, once 1 leaf suffered from rice leaffolder damage, the entire plant was considered to be sensitive; otherwise, the plants were regarded as resistant. For R_2 homozygous lines, 20 plants were randomly selected to investigate the number of total leaves and the number of damaged leaves.

RESULTS

Difference in codon usage bias between rice and *Bt*

An analysis of codon usage information downloaded from the condon usage database showed that rice and *Bt* differed greatly in the GC content of the CDSs. First, the total GC content of rice was up to 55.26%, which was much higher than the GC content of *Bt* (35.46%). The difference between the GC content of the species is 19.80%. Second, the GC content in 3 different positions of codons also showed a significant difference between the 2 species. The third position had the largest difference in GC content between rice (61.61%) and *Bt* (24.55%), and the difference reached 37.06%. The first position had the second largest difference in GC content between rice (58.19%) and *Bt* (45.82%). The second position showed the smallest difference between rice (45.97%) and *Bt* (36.02%); even so, this difference was still higher than 9% (Figure 2).

The usage frequency of various codons in the 2 species was calculated in Table 1, and the results were interpreted as follows. First, the usage of synonymous codons did not appear to be random because certain codons were used more frequently than others. For example, in rice, CUC had the highest usage frequency (28.51%) among all synonymous codons of leucine, whereas UUA had the lowest usage frequency, only 6.79%; the difference between the values reached 4.2 fold. Second, between rice and *Bt*, with the exception of tryptophan and

methionine, which have no synonymous codons, the codons with the highest usage frequency (preferred codons), including termination codons, varied. The preferred codons in rice were TTC, CTC, TCC, AGC, TAC, TGC, CCG, CAC, CAG, CGC, ATC, ACC, AAC, AAG, GTG, GCC, GAC, GAG, GGC, and TGA. In turn, the preferred codons in *Bt* were TTT, TTA, TTG, AGT, TAT, TGT, CCA, CAT, CAA, AGA, ATT, ACA, AAT, AAA, GTA, GCA, GAT, GAA, GGA, and TAA.

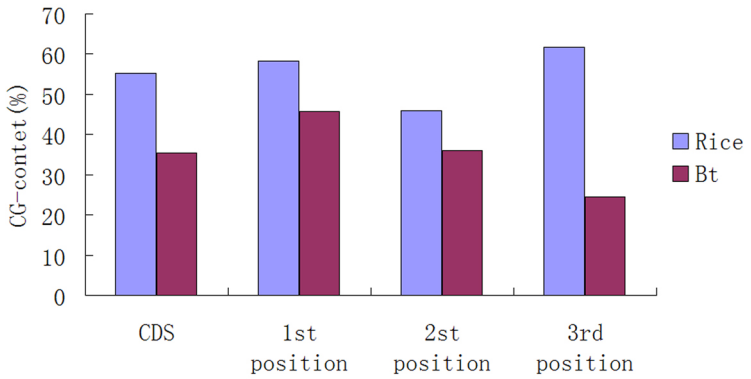

Figure 2. Differences in GC content of coding DNA sequences (CDSs) and individual codon positions between rice and *Bacillus thuringiensis* (*Bt*).

According to the analysis mentioned above, there was a large difference between codon usage by rice and *Bt*. Rice preferred to use codons with high GC content, whereas *Bt* preferred codons with low GC content.

Design of *mCry1Ab* and sequence analysis

Amino acids of the Cry1Ab protein were individually replaced with the rice-preferred codons in Table 1, except for serine and arginine, for which 2 codons were used because the 2 codons had similar usage frequency among synonymous codons. Additionally, the rice-preferred termination codon TGA was added to the end of the new DNA sequence to form a new gene called *mCry1Ab*.

A comparison of the DNA sequences between *Cry1Ab* and *mCry1Ab* resulted in the following conclusions. First, a total of 565 bases in *mCry1Ab* were modified compared with *Cry1Ab*, accounting for 30.62% of the total length. Second, *Cry1Ab* used 58 of the 61 codons for 20 amino acids, whereas *mCry1Ab* only employed 22 codons, and 39 codons were not used in this gene. Third, there was a large difference in GC content between the 2 genes. The GC content of *Cry1Ab* was 37%, whereas that of *mCry1Ab* was up to 65%, reaching the maximum value. Fourth, *Cry1Ab* had 33 fragments with 6 or more consecutive adenine (A) or thymine (T) nucleotides, 2 plant transcription termination signal (poly A) sequences of AATAATT, and 11 sequences of ATTTA that could cause mRNA instability in plants, whereas *mCry1Ab* had no AATAATT or ATTTA sequences and did not contain consecutive strings of 3 or more A+T nucleotides. However, after the GC content increased, consecutive G+C strings in *mCry1Ab* occurred frequently. There were only 2 sites of consecutive G+C strings (>5 bases) in the

Table 1. Usage frequency of codons in rice and *Bacillus thuringiensis* (*Bt*).

Codon	Amino acid	Rice		*Bt*	
		Codon No.	Percentage (%)	Codon No.	Percentage (%)
UUU	F	446,063	36.88	6833	75.77*
UUC		763,386	63.12*	2185	24.23
UUA	L	209,911	6.79	9062	35.77*
UUG		500,913	16.21	9062	35.77*
CUU		517,919	16.76	3100	12.24
CUC		880,959	28.51*	730	2.88
CUA		263,871	8.54	2499	9.87
CUG		716,525	23.19	878	3.47
UCU	S	433,684	16.17	3999	26.62
UCC		557,258	20.77*	1170	7.79
UCA		424,426	15.82	3365	22.40
UCG		420,825	15.69	1028	6.84
AGU		300,879	11.22	4136	27.53*
AGC		545,529	20.34*	1327	8.83
UAU	Y	339,638	39.65	7865	81.25*
UAC		517,042	60.35*	1815	18.75
UGU	C	211,609	33.35	1458	71.09*
UGC		422,851	66.65*	593	28.91
UGG	W	472,543	100.00	2469	100.00
CCU	P	463,459	23.47	2587	32.54
CCC		411,848	20.86	347	4.36
CCA		486,283	24.63	3720	46.79*
CCG		613,159	31.05*	1296	16.30
CAU	H	385,174	44.99	3198	81.92*
CAC		470,960	55.01*	706	18.08
CAA	Q	460,234	39.35	7318	81.17*
CAG		709,469	60.65*	1698	18.83
CGU	R	244,821	10.30	2050	22.52
CGC		550,575	23.17*	586	6.44
CGA		219,662	9.24	1433	15.74
CGG		458,602	19.30	333	3.66
AGA		358,226	15.07	3678	40.40*
AGG		544,515	22.91*	1023	11.24
AUU	I	483,941	33.46	7965	57.04*
AUC		662,207	45.79*	2118	15.17
AUA		300,085	20.75	3881	27.79
AUG	M	312,432	100.00	3856	100.00
ACU	T	363,451	21.97	4271	30.28
ACC		508,156	30.71*	1252	8.88
ACA		394,803	23.86	6399	45.37*
ACG		388,036	23.45	2183	15.48
AAU	N	515,761	44.90	11,652	77.54*
AAC		633,017	55.10*	3,75	22.46
AAA	K	544,476	33.08	9828	76.45*
AAG		1,101,342	66.92*	3027	23.55
GUU	V	529,509	23.27	4191	31.29
GUC		685,971	30.14	1180	8.81
GUA		231,761	10.18	5673	42.35*
GUG		828,681	36.41*	2351	17.55
GCU	A	667,854	20.75	3979	32.97
GCC		1,050,723	32.65*	996	8.25
GCA		591,267	18.37	5169	42.83*
GCG		908,634	28.23	1924	15.94
GAU	D	863,983	47.38	9112	82.45*
GAC		959,498	52.62*	1939	17.55
GAA	E	738,891	35.96	10,790	74.20*
GAG		1,315,826	64.04*	3752	25.80
GGU	G	505,609	19.17	4593	34.60
GGC		1,005,701	38.14*	1385	10.43
GGA		542,264	20.56	5229	39.40*
GGG		583,357	22.12	2066	15.57
UAA	STOP	22,360	24.24	331	66.73*
UAG		28,508	30.91	114	22.98
UGA		41,361	44.85*	51	10.28
Total		34,132,283		211,286	

*Preferred codons with the highest usage frequency.

 ©FUNPEC-RP www.funpecrp.com.br

original gene, which were located at 556 bp with 6 bases (GCCGCG) and 986 bp with 7 bases (GGGGCC). However, in *mCry1Ab*, there were 23 consecutive G+C strings (>5 bases) broadly spanning the entire gene (Figure 3).

Figure 3. Sequence alignment of *Cry1Ab* and *mCry1Ab*. AATAATT, highlighted in green, plant transcription termination signal; ATTTA, highlighted in magenta, unstable sequence of plant mRNA. The underlined sequences are the consecutive G+C strings (>5 bases).

FMCry1Ab, which was modified by Perlak et al. (1991), had the same protein sequence as *mCry1Ab* and used 55 codons, but it had a GC content of 49% (Table 2), which is far lower than that of *mCry1Ab*. Consecutive G+C strings (>4 bases) were avoided when designing *FMCry1A*b although they were broadly distributed in *mCry1Ab*. The results suggested that *mCry1Ab* was vastly different from both wild-type *Cry1Ab* and *FMCry1A*b. Basic local alignment search tool results from the GenBank database (http://blast.ncbi.nlm.nih.gov/) showed that there was no previous record of the *mCry1Ab* sequence.

Table 2. Base components in *mCry1Ab*, *Cry1Ab*, and *FMCry1Ab*.

Gene	Codon No.	GC content (%)
Cry1Ab	58	37
FMCry1Ab	55	49
mCry1Ab	22	65

Transformation and PCR assay of R_0 transgenic plants

To determine whether *mCry1Ab* was functional in rice, we fused *mCry1Ab* to the constitutive promoter *CaMV35S* and the tissue-specific promoter *OsTSP I* (Figure 1) and transferred them into rice via *Agrobacterium*-mediated transformation to obtain 22 p*35s::mCry1Ab* transgenic seedlings and 47 p*OsTSPI::mCry1Ab* transgenic seedlings (Figure 4a). A total of 69 independent transgenic plants were detected by PCR analysis, and the target band of 158 bp could be amplified from all of the plants (Figure 4b).

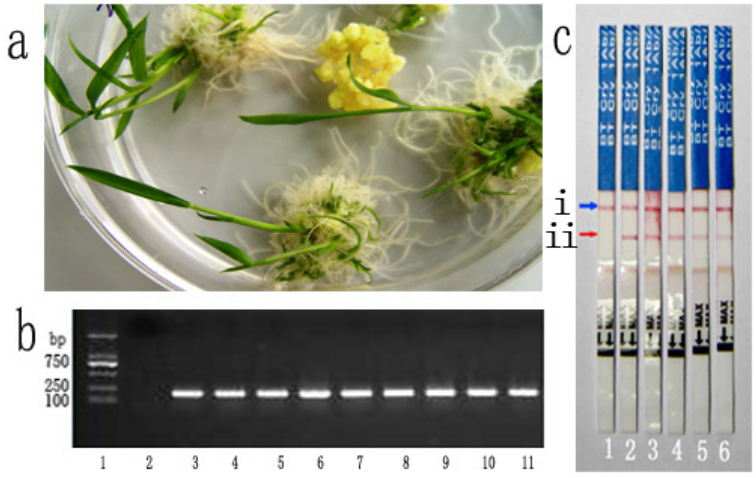

Figure 4. Transgenic seedlings, polymerase chain reaction (PCR) analysis, and Cry1Ab/Cry1Ac test strip detection. **a.** Regenerated plantlets. **b.** PCR assay of transformants (*lane 1* = ladder marker; *lane 2* = negative control; *lane 3* = positive plasmid DNA; *lanes 4-11* = transformed plants). **c.** Cry1Ab/Cry1Ac strip assay (i = quality control line; ii = detection line; *lane 1* = negative control; *lanes 2-6* = leaves of transformed plants).

Detection of Cry1Ab protein expressed in transgenic plants

Transgenic rice leaves were detected using Cry1Ab/Cry1Ac strips (Figure 4c). Seventeen of the 22 p*35S::mCry1Ab*-transgenic seedlings resulted in positive plants, yielding a positive rate of 77.3%; 39 positive plants were obtained from the p*OsTSPI::mCry1Ab*-transgenic seedlings, yielding a positive rate of 88.6%.

The toxic protein content of fresh leaves (tillering stage) from positive plants was measured by ELISA. The mean Cry1Ab content in the p*35S::mCry1Ab* and p*OsTSPI::mCry1Ab* transgenic plants was 6.02 ± 1.05 and 8.43 ± 2.6 µg/g fresh leaf (means ± SD, N = 17 and 39), respectively.

Bioassay of *Cry1Ab* plants in the laboratory

To test whether the positive plants had insecticidal activity toward rice leaffolder, we conducted a laboratory bioassay by infesting R_0 leaves with rice leaffolder. The results indicated that the positive plants showed a high level of insecticidal activity toward rice leaffolder.

Larvae of leaffolder either refused or showed reluctance to feed on explants of transgenic plants as evident from less feeding damage on the infested leaf sections (Figure 5a 3-4) compared with heavily damaged control explants (Figure 5a 1). Larval growth and development were also inhibited after they were exposed to explants from the positive plants, although a few remained alive after 4 days of infestation (Figure 5b 2). In contrast, feeding damage to the leaves of pCAMBIA1305.1-transgenic plants (negative control) was severe during this period, and the leaffolder larvae grew and developed normally (Figure 5b 1). The comparative response of leaffolder larvae to the positive transgenic plants and the negative control plants shows that *mCry1Ab* that was driven by the *CaMV35S* and *OsTSP I* promoters in rice was effective against the lepidopteran insect.

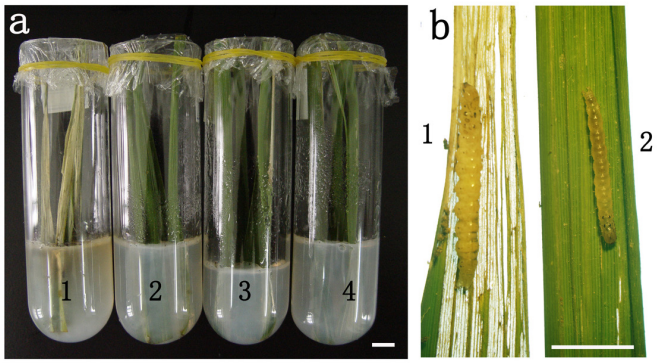

Figure 5. Bioassay for insecticidal activity of Cry1Ab in transgenic rice lines against rice leaffolder. Damage to explants and insect performance were observed and photographed 4 days after infestation. **a.** Leaf explants from plants of the pCAMBIA1305.1-transgenic line (Tube 1), p*35S::mCry1Ab*-transgenic line (Tube 3), and p*OsTSPI::mCry1Ab*-transgenic line (Tube 4) that were artificially infested with second to third instar leaffolder larvae in test tubes. Tube 2 contained leaf explants from plants of the pCAMBIA1305.1-transgenic line without leaffolder larvae. **b.** Feeding damage on leaves of the pCAMBIA1305.1-transgenic line (Leaf 1) and p*OsTSPI::mCry1Ab*-transgenic line (Leaf 2) caused by leaffolder larvae. Scale bar = 1 cm.

Field resistance of transgenic rice plants to rice leaffolder

To assess the resistance of the transgenic *mCry1Ab* plants in the field, 69 R_0 transgenic plants and 30 pCAMBIA1305.1-transgenic plants (negative control) were planted in a net house and surrounded with wild-type rice to induce outbreaks of rice leaffolder under natural conditions. During the entire growing period, pesticide was prohibited. Fifteen days after flowering, every plant of the negative control and wild-type rice suffered severe damage such that the leaf blade was folded and glued with silk strands and had longitudinal white and transparent streaks due to rice leaffolder feeding, whereas most of the transgenic *mCry1Ab* lines did not show any symptoms. For the p*35S*::*mCry1Ab*-transgenic lines, 18 of 22 showed no damage; of these, 17 were positive in the Cry1Ab/Cry1Ac strip assay, and 4 lines that were heavily damaged by rice leaffolder were negative in the Cry1Ab/Cry1Ac strip assay. p*OsTSPI*::*mCry1Ab*-transgenic lines showed similar results to those of the p*35S*::*mCry1Ab*-transgenic lines. Thirty-eight of 47 p*OsTSPI*::*mCry1Ab*-transgenic lines showed no damage; of these, 36 were positive in the Cry1Ab/Cry1Ac strip assay. However, 3 negative lines in the Cry1Ab/Cry1Ac assay were found to be intact, possibly because they were surrounded by

positive lines and unexposed to insect larvae. Additionally, there were 3 strip-positive lines (p*35S*::*mCry1Ab*) whose leaf blades were folded and glued with silk strands by leaffolder but remained green without bite streaks. We also categorized these as susceptible plants.

To confirm the relationship between *mCry1Ab* expression and insect resistance, the chi-squared test for independence was used, and the result showed that insect resistance was significantly correlated with *mCry1Ab* expression (χ^2 = 11.16 and 15.77 >> $\chi^2_{0.01,1}$ = 6.63) (Table 3), suggesting that *mCry1Ab* expression endows rice with insect resistance.

Table 3. Correlation between the strip assay and insect resistance of transgenic plants in the field.

Vector	Strip assay	Insect damage	χ^2 value	P value
p*35S::mCry1Ab*	Positive plants: 17	Damaged by insects: 0 Not damaged by insects: 17	11.16	0.9992
	Negative plants: 5	Damaged by insects: 4 Not damaged by insects: 1		
p*OsTSP I::mCry1Ab*	Positive plants: 39	Damaged by insects: 3 Not damaged by insects: 36	15.77	0.9999
	Negative plants: 8	Damaged by insects: 6 Not damaged by insects: 2		

Thirty R_1 transgenic plant families derived from each R_0 transgenic plant were grown in a net house in 2010 to investigate insect resistance. No insecticide was used during the rice growth period. In theory, the segregation ratio of resistance *vs* sensitivity in every family should be 3:1 or higher. However, abnormal segregations (1:1 or less than 1:1) in R_1 transgenic progeny were found in this study, and similar results were also observed in previous studies (Husnain et al., 2002; Wang et al., 2002; Chen et al., 2005). In this study, approximately 25% of families segregated in the ratio 3:1 (resistance *vs* sensitivity), and the resistant plants grew normally without folded leaves and deadhearts (Figure 6a). However, the sensitive plants and control plants were heavily damaged by lepidopteran insects, embodying transparent streaks on folded blades and deadhearts (Figure 6b). These resistant families were chosen to develop homozygous lines with a single copy of the inserted transgene.

Figure 6. Field performance of the R_1 and R_2 generations of transgenic mCry1Ab lines under natural conditions. **a.** Positive plants isolated from R_1 transgenic families resisted lepidopteron insects. **b.** Wanjing97 control plants were damaged seriously by the rice leaffolders. **c.** Damage of R_2 transgenic plant leaves and wild-type leaves. Leaf 1 was a leaf of Wanjing97 that was folded longitudinally by leaffolders that scraped the green mesophyll on the inside blade resulting in linear, pale, white stripe damage, and the general vigor and photosynthetic ability of the infested leaf was greatly reduced. Leaf 2 was a leaf of the R_2-231 line that was immune to leaffolders, showing no symptoms. Leaf 3 was a leaf of the R_2-231 line showing no whitish appearance, although it was folded and glued with silk strands by leaffolders. Scale bar = 5 cm.

To obtain homozygous transgenic plants, the mature seeds from the selected R_1 transgenic plants were collected separately and germinated in Petri dishes with sterilized H_2O at 30°C. After 4 days, sterilized H_2O was replaced with 50 mg/L hygromycin solution. Five days later, negative seedlings withered, but the positive seedlings grew normally. Because the survival ratios of the homozygous transgenic plants, heterozygous plants, and negative plants were 100, 75, and 0%, respectively, the selected families containing a single insertion site could be confirmed.

Two homozygous lines, R_2-17 (p*35S::mCry1Ab*-transgenic line) and R_2-231 (p*OsTSPI::mCry1Ab*-transgenic line), which had no significant phenotypic changes compared with the wild-type line, were selected to investigate the degree of insect damage under natural conditions. Twenty plants per line were chosen to count total leaves and folded leaves. The resistance performance of the 2 *mCry1Ab* transgenic lines was very similar and was significantly superior to that of the original Wanjing97 recipient (Table 4). During the entire growing period, Wangjing97 was damaged seriously by the rice leaffolders (Figure 6c 1), and the number of damaged leaves per tiller (NDLPT) was so high, 5.45, that every tiller had damaged leaves. In contrast, most tillers from the R_2-17 and R_2-231 lines did not show any symptoms, and the NDLPTs of the 2 lines were 0.20 and 0.03, respectively. The rate of folded leaves was as high as 64.19% in Wanjing97, whereas the occurrence of folded leaves was only 2.62% in R_2-17 and 0.47% in R_2-231. The resistance of R_2-231 seemed to be superior to that of R_2-17, most likely indicating that the Cry1Ab protein content in R_2-231 was twice greater than that of R_2-17 (Table 4).

Table 4. Folded leaf ratio as a measure of field resistance in two *mCry1Ab* transgenic lines (natural infestation at the tillering stage).

Plant line	Cry1Ab protein content (μg/g)	Tillers	Total leaves	Folded leaves	Folded leaves/Tillers	Folded leaves/ Total leaves (%)
Wangjing97	0	78	662	425	5.45	64.19
R_2-17	4.2 ± 0.54*	83	649	17	0.20	2.62
R_2-231	8.3 ± 0.79*	87	633	3	0.03	0.47

*Means ± SD, N = 5.

In fact, R_2-17 and R_2-231 were not 100% immune to leaffolders: a few leaves were also hurt by the insect (Table 4). However, the degree of harm to these leaves was different from that of the negative control. These leaves were only folded and glued by silk strands, and the feeding area was less than 5% (Figure 6c 3); however, the feeding area of the control plant leaves usually exceeded 30% of the total area (Figure 6c 1). These results indicated that the leaffolders could be well controlled in the *mCry1Ab*-transgenic rice.

Transgenic plants with *mCry1Ab* driven by constitutive and tissue-specific promoters were found to be highly resistant to rice leaffolders in the R_0, R_1, and R_2 generations in the field, indicating that the modified *mCry1Ab* gene could be expressed and inherited stably in rice plants.

DISCUSSION

Transgenic technology can break reproductive isolation to realize gene transfer among different species, and rice is a potential plant that can be used as a recombinant protein expres-

sion system (Liu et al., 2011; Kim et al., 2012; An et al., 2013; Kuo et al., 2013). However, after a high-expression gene in one species it is transferred to another species, the expression is often reduced greatly, particularly for species with a large expression difference, leading to the loss of the gene's function (Gustafsson et al., 2004). Furthermore, the use of different promoters, fusion proteins, and leader sequences cannot significantly increase exogenous gene expression (Perlak et al., 1991).

One of the main reasons for the low-level expression is the codon usage bias in different species. Wang et al. (2004) analyzed the entire sequence of the rice genome and demonstrated that there is a positive correlation between the usage frequency of codons and the quantity of the corresponding tRNA genes in the genome. If species have different biased codons, their tRNA number or content could also vary largely (Ikemura, 1985), and disfavored codons are translated less rapidly (Dix and Thompson, 1989). In this study, we provided a set of the rice-preferred codons for each amino acid - e.g., CUC for leucine-to make the gene modification method more species specific than using G or C to randomly replace A and T at the third position of triplet codons.

Another main reason for the low-level expression is intron splicing. Unlike microbes, intron splicing in plants is essential for generating mature mRNA (Gutiérrez et al., 1999). However, this splicing mechanism may recognize exogenous genes as introns and cut them, causing gene silencing. In plants, exons usually contain a high GC content, whereas introns have a high AT content (Goodall and Filipowicz, 1989). The abundant AU sequences in precursor nuclear messenger RNA, specifically AU islands consisting of 4-7 nt, would be bound by nuclear proteins or reduce secondary structure formation, forming splicing sites nearby (Goodall and Filipowicz, 1990; Brown and Simpson, 1998). *Cry1Ab* contained 33 consecutive A+T strings (> 6 bases), including 2 plant transcription termination signal (poly A) sequences of AATAATT and 11 sequences of ATTTA that may lead to unstable mRNA in plants, most likely making it difficult to form complete mRNAs when the gene is transformed into plant cells. Perlak et al. (1991) avoided consecutive strings of 5 or more A+T nucleotides in *FMCry1Ab*. We did not specifically avoid the consecutive A+T sequences in *mCry1Ab*, but *mCry1Ab* contained neither AATAATT nor ATTTA sequences, and it even lacked consecutive strings of 3 or more A+T nucleotides mainly because all rice-preferred codons tended to include the codon with the highest GC content among synonymous codons and increased the GC content of *mCry1Ab*.

However, high GC content caused the emergence of potential methylation signals (Bender, 2004) in *mCry1Ab*. The GC content of *mCry1Ab* was 65%, reaching the maximum. The high GC content caused the number of continuous G+C ($N \geq 6$) sequences to increase from 2 in the original *Cry1Ab* to 23 in *mCry1Ab*. Murray et al. reported that CG should be avoided in the genetic engineering of plants and other eukaryotes and that G nucleotides should be avoided in the third position of the triplet codons for threonine, proline, alanine, and serine because the second base was C, and CG may be involved in methylation and subsequently cause gene silencing (Murray et al., 1989). However, we chose CCG as the codon of proline when designing *mCry1Ab*. Our results showed that *mCry1Ab* could be expressed and inherited stably in subsequent generations. Furthermore, we searched the Rice Oligonucleotide Array Database (http://www.ricearray.org/expression/highly.shtml) for the top 50 highly expressed genes in whole plant and found that the GC content of the endogenous gene CDSs ranged from 37.65 to 73.21% (1 gene < 40%, 9 genes = 40-50%, 13 genes = 50-60%, 21 genes = 60-70%, and 6 genes > 70%), and the average was 60.41%, which was higher than the mean GC content

of 55% for the CDSs of all the rice genes. Thus, high GC content may not be the limiting factor that affects efficient gene expression in rice.

mCry1Ab included 22 of the 61 codons, and the other 39 codons were not used in this gene. The normal expression and stable inheritance of *mCry1Ab* indicated that it was not necessary to use all codons for gene modification and that simply choosing 1 or 2 biased codons could ensure the expression of exogenous genes in rice.

Generally, plants showing a superior resistance effect need to contain over 3.0 μg/g toxin in the fresh leaf (Cai et al., 2007). In this study, the expression quantity of *mCry1A* driven only by the *CaMV35S* promoter (without a castor been intron) could reach 6 μg/g toxin in fresh leaf, which is 2-fold greater than the superior resistance threshold; thus, *mCry1A* would be valuable for transgenic rice breeding.

ACKNOWLEDGMENTS

We greatly thank Dr. Yuping Lu (Beijing DBN Technology Group Co., Ltd., China) for providing help with designing *mCry1Ab*. Financial support was from the National Natural Science Foundation of China (#31201276), the Special Research Projects for Developing Transgenic Plants of China (#2013ZX08010002-009), the National Science Foundation of Anhui Province (#1408085MC55, #1408085QC53), and the Youth Foundations of President of the Anhui Agricultural Academy of Sciences (#11B0108, #13B0101, #13B0102, #13B0113, #13B0114 and #13C0101).

REFERENCES

An N, Ou J, Jiang D, Zhang L, et al. (2013). Expression of a functional recombinant human basic fibroblast growth factor from transgenic rice seeds. *Int. J. Mol. Sci.* 14: 3556-3567.

Barton KA, Whiteley HR and Yang NS (1987). *Bacillus thuringiensis* section sign-endotoxin expressed in transgenic *Nicotiana tabacum* provides resistance to Lepidopteran insects. *Plant Physiol.* 85: 1103-1109.

Bender J (2004). DNA methylation and epigenetics. *Annu. Rev. Plant Biol.* 55: 41-68.

Brown JW and Simpson CG (1998). Splice site selection in plant pre-mRNA splicing. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 49: 77-95.

Cai M, Wei J, Li X, Xu C, et al. (2007). A rice promoter containing both novel positive and negative cis-elements for regulation of green tissue-specific gene expression in transgenic plants. *Plant Biotechnol. J.* 5: 664-674.

Chen H, Tang W, Xu C, Li X, et al. (2005). Transgenic indica rice plants harboring a synthetic *cry2A** gene of *Bacillus thuringiensis* exhibit enhanced resistance against lepidopteran rice pests. *Theor. Appl. Genet.* 111: 1330-1337.

Dix DB and Thompson RC (1989). Codon choice and gene expression: synonymous codons differ in translational accuracy. *Proc. Natl. Acad. Sci. U. S. A.* 86: 6888-6892.

Duan Y, Zhai C, Li H, Li J, et al. (2012). An efficient and high-throughput protocol for *Agrobacterium*-mediated transformation based on phosphomannose isomerase positive selection in Japonica rice (*Oryza sativa* L.). *Plant Cell Rep.* 31: 1611-1624.

Ferré J and Van Rie J (2002). Biochemistry and genetics of insect resistance to *Bacillus thuringiensis*. *Annu. Rev. Entomol.* 47: 501-533.

Fujimoto H, Itoh K, Yamamoto M, Kyozuka J, et al. (1993). Insect resistant rice generated by introduction of a modified delta-endotoxin gene of *Bacillus thuringiensis*. *Biotechnology* 11: 1151-1155.

García-Robles I, Ochoa-Campuzano C, Fernández-Crespo E, Camañes G, et al. (2013). Combining hexanoic acid plant priming with *Bacillus thuringiensis* insecticidal activity against Colorado Potato Beetle. *Int. J. Mol. Sci.* 14: 12138-12156.

Goodall GJ and Filipowicz W (1989). The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. *Cell* 58: 473-483.

Goodall GJ and Filipowicz W (1990). The minimum functional length of pre-mRNA introns in monocots and dicots. *Plant*

Mol. Biol. 14: 727-733.
Gustafsson C, Govindarajan S and Minshull J (2004). Codon bias and heterologous protein expression. *Trends Biotechnol.* 22: 346-353.
Gutiérrez RA, MacIntosh GC and Green PJ (1999). Current perspectives on mRNA stability in plants: multiple levels and mechanisms of control. *Trends Plant Sci.* 4: 429-438.
Husnain T, Asad J, Maqbool SB and Datta SK (2002). Variability in expression of insecticidal *cry1Ab* gene in indica Basmati rice. *Euphytica* 128: 121-128.
Ikemura T (1985). Codon usage and tRNA content in unicellular and multicellular organisms. *Mol. Biol. Evol.* 2: 13-34.
Kim YM, Lee JY, Lee T and Lee YH (2012). The suppression of the glutelin storage protein gene in transgenic rice seeds results in a higher yield of recombinant protein. *Plant Biotechnol. Rep.* 6: 347-353.
Kuo YC, Tan CC, Ku JT, Hsu WC, et al. (2013). Improving pharmaceutical protein production in *Oryza sativa*. *Int. J. Mol. Sci.* 14: 8719-8739.
Li H, Li J, Cong XH, Duan YB, et al. (2013). A high-throughput, high-quality plant genomic DNA extraction protocol. *Genet. Mol. Res.* 12: 4526-4539.
Liu F, Zhao YY and Zheng JG (2011). High-level expression of modified gene encoding human adiponectin in transgenic rice. *Biol. Res.* 44: 369-375.
Murray EE, Lotzer J and Eberle M (1989). Codon usage in plant genes. *Nucleic Acids Res.* 17: 477-498.
Nakamura Y, Gojobori T and Ikemura T (2000). Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res.* 28: 292.
Park SY, Lee SM, Lee JH and Ko HS (2012). Compositional comparative analysis between insect-resistant rice (*Oryza sativa* L.) with a synthetic cry1Ac gene and its non-transgenic counterpart. *Plant Biotechnol. Rep.* 6: 29-37.
Perlak FJ, Fuchs RL, Dean DA, McPherson SL, et al. (1991). Modification of the coding sequence enhances plant expression of insect control protein genes. *Proc. Natl. Acad. Sci. U. S. A.* 88: 3324-3328.
Qiu C, Sangha JS, Song F, Zhou Z, et al. (2010). Production of marker-free transgenic rice expressing tissue-specific Bt gene. *Plant Cell Rep.* 29: 1097-1107.
Siegel JP (2001). The mammalian safety of *Bacillus thuringiensis*-based insecticides. *J. Invertebr. Pathol.* 77: 13-21.
Tang W, Chen H, Xu C and Li X (2006). Development of insect-resistant transgenic indica rice with a synthetic *cry1C** gene. *Mol. Breed.* 18: 1-10.
Tu J, Datta K, Alam MF, Fan Y, et al. (1998). Expression and function of a hybrid *Bt* toxin gene in transgenic rice conferring resistance to insect pest. *Plant Biotechnol.* 15: 195-203.
Vaeck M, Reynaerts A, Höfte H and Jansens S (1987). Transgenic plants protected from insect attack. *Nature* 328: 33-37.
Wang ZG, Shu QY, Ye GY, Cui HR, et al. (2002). Genetic analysis of resistance of *Bt* rice to stripe stem borer (*Chilo suppressalis*). *Euphytica* 123: 379-386.
Wang XY, Shi XL and Hao BL (2004). The tRNA and rRNA genes in the *Oryza sativa* genome. *Yi Chuan Xue Bao* 31: 871-877.
Zaidi MA, Ye G, Yao H, You TH, et al. (2009). Transgenic rice plants expressing a modified *cry1Ca1* gene are resistant to *Spodoptera litura* and *Chilo suppressalis*. *Mol. Biotechnol.* 43: 232-242.