Frequency of Mutations of the Transgene, which might Result in the Loss of the Glyphosate-Tolerant Phenotype, was Lowered in Roundup Ready® Soybeans

Takeshi Ogasawara,[a] Yukie Chikagawa,[a] Fumihiro Arakawa,[b] Asami Nozaki,[a] Yoshio Itoh,[a] Kazuo Sasaki,[c] Hironori Umetsu,[c] Takahiro Watanabe,[d] Hiroshi Akiyama,[d] Tamio Maitani,[d] Masatake Toyoda,[e] Hiroshi Kamada,[f] Yukihiro Goda,[d] and Yoshihiro Ozeki*,[a]

[a]*Department of Biotechnology, Faculty of Technology, Tokyo University of Agriculture and Technology, 2–24–16 Naka-cho, Koganei, Tokyo 184–8588, Japan,* [b]*San-Ei Gen F.F.I., Inc., 1–1–11 Sanwa-cho, Toyonaka, Osaka 561–8588, Japan,* [c]*Department of Clinical Pharmacy, Faculty of Pharmaceutical Sciences, Aomori University, 2–3–1 Kobata, Aomori 030–0943, Japan,* [d]*National Institute of Health Sciences, 1–18–1 Kamiyoga, Setagaya, Tokyo 158–8501, Japan,* [e]*Department of Food and Health Sciences, Faculty of Human Life Sciences, Jissen Women's University, 4–1–1 Osakaue, Hino, Tokyo 191–8510, Japan, and* [f]*Institute of Biological Sciences, University of Tsukuba, 1–1–1 Tennoudai, Tsukuba, Ibaraki 305–8572, Japan*

(Received June 30, 2004; Accepted December 9, 2004; Published online December 13, 2004)

Polymerase chain reaction (PCR) primers were prepared to amplify the DNA fragment between the genomic DNA sequence adjacent to the 5′-integration site of Roundup Ready® (RR) soybeans neighboring the transgene and the parts of the coding region of the transgene, together with the primer set for the internal host gene, the α' subunit of β-conglycinin storage protein gene (*Cong* gene). Using the primers for the transgene and *Cong* gene, the DNA fragments were amplified from the individual genomic DNAs prepared from 72 samples of RR soybean isolated from imported soybean seeds labeled "not segregated." Although the frequency of alterations of the nucleotide sequences in both the transgene and *Cong* gene were almost the same, the mutations that caused alterations to the amino acid sequence were more highly repressed in the transgene than in the *Cong* gene. In the nucleotide sequence upstream of the coding region of the transgene, the number of alterations of the nucleotide in the proximal promoter region was smaller than that in the further upstream region, suggesting that the mutants missing or being weak glyphosate-tolerance by an alteration of the critical nucleotide sequences in the promoter or coding region might be discarded artificially. It is supposed that the selective bias on the transgene might be extremely high, which indicates that the nucleotide sequence of the transgene might be stable and maintained in inbred RR soybean lines.

Key words —— genetically modified crop, genetically modified organism, mutation, Roundup Ready soybean, soybean, transgene

INTRODUCTION

It has been suspected that the transgene might be a "hot spot" for mutations in the genetically modified (GM) organisms, but current research into commercially important GM crops has not yet indicated if this suspicion is correct. It is true that some reports showed instability of the transgenes caused by rearrangements of transforming DNA in regenerated calli and plantlets just after transformation.[1,2] However, such instability should be prevented to establish commercial GM cultivars. The transgenes should also be inherited correctly and in a stable and predictable manner over successive generations. Field trial experiments showed a stable inheritance of the phenotype in the progeny of GM crops, and in other experiments the stability of the transgenes in individual plants of the progeny of GM crops were found using a Southern blot analysis.[3–5] These results suggested that recombination or rearrangement of the transgenes required to drastically alter their structure did not occur, or occurred to only a limited extent, in the progeny of GM crops. However, the alteration, substitution or deletion of transgenes at the nucleotide sequence level has not yet been investigated in individual plants of the progeny of GM crops.

We isolated and identified the genomic sequence adjacent to the 5′-integration site of the transgene of Roundup Ready® (RR) soybeans.[6] Using this sequence, we prepared the primer pair amplified for

*To whom correspondence should be addressed: Department of Biotechnology, Faculty of Technology, Tokyo University of Agriculture and Technology, 2–24–16 Naka-cho, Koganei, Tokyo 184–8588, Japan. Tel. & Fax: +81-42-388-7239; E-mail: ozeky@cc.tuat.ac.jp

the proximal insertion region, the cauliflower mosaic virus 35S promoter (designated as 35S promoter) and part of the petunia chloroplast transit peptide (*CTP*) gene in the RR soybean. We also constructed the primers for the promoter and coding regions of the soybean α' subunit of β conglycinin storage protein genes (designated *Cong* genes below)[7,8] as controls of the internal host gene. β conglycinin is a major component of seed storage proteins in the soybean (up to 30% of total seed protein) and is composed of three subunits: α, α' and β. Their high expression ensures the synthesis and accumulation of storage protein and affects the nutritive quality of seed crops. As such, these genes might have been maintained by breeders. Using the primers for the transgene and internal gene, DNA fragments were amplified with the genomic DNA prepared from 72 individual RR soybeans isolated from batches of imported soybeans with an "unsegregated" label from U.S.A. that had been sampled by the quarantine station of the Ministry of Health, Welfare and Labor (MHWL). The nucleotide sequences of the amplified DNA fragments using these primer pairs were determined, and the mutation rates of the nucleotide and deduced amino acid sequences in the transgene were analyzed in comparison with the internal *Cong* genes.

MATERIALS AND METHODS

Preparation of Genomic DNA from Soybean and Identification of the RR Soybean —— Soybean (*Glycine max*) seeds imported from U.S.A. labeled as "not segregated" and sampled by the quarantine station of the MHLW were used in this study. The genomic DNAs were prepared from individual seeds of several sampled batches using a modified CTAB method.[9] In order to identify which genomic DNA preparations were derived from individual RR soybean seeds, the genomic DNAs were subjected to a qualitative PCR using the primers for the nucleotide sequences of the 35S promoter and 5-enol-pyruvylshikimate-3 phosphate synthase (*EPSPS*) gene.[9]

PCR Amplification of the Genomic DNA Prepared from RR Soybeans —— In order to isolate the DNA fragments that corresponded to the proximal insertion region-35S promoter-*CTP* (accession number I07318), the primer pair of CHM1, 5′-ACTGAAAAATTCAGAACCTTGTGC-3′, and CTP-R, 5′-TTCAAAACCAACATAGAATTTGCTG-3′, was prepared. The other pair, C-1, 5′-GCTGATCAGGATCGCCGCGTC-3′, and C-2R, 5′-GAATGGACGTGGCTGCTCACC-3′, was prepared to isolate the fragments that corresponded to the proximal promoter and coding region of the *Cong* gene (accession number M13759). A DNA extract (4 μl) of 100 ng prepared above was added to a reaction mixture that consisted of 1 × LA-PCR buffer (TaKaRa Bio Inc., Shiga, Japan), 2.5 mM $MgCl_2$, 0.25 mM dNTP and 0.5 μM of the primers. The mixture was made up to a total volume of 20 μl with water, and the equal volume of mineral oil was added on the mixture. The tubes of the reaction mixture were set on the PCR machine (MiniCycler™, MJ Japan Ltd., Tokyo, Japan) and incubated at 98°C for 1 min, after which the temperature was lowered to 92°C. The cap of the tube was opened and 0.25 units (1 μl) of LA *Taq* DNA polymerase (TaKaRa Bio Inc.) were added to the bottom of the reaction mixture using a pipette. The cap was closed and PCR proceeded for 35 cycles of denaturing at 92°C for 30 sec, annealing at 52°C for 45 sec and extension at 72°C for 1 min. Following this, the mixture was incubated at 72°C for 10 min, and then cooled at 4°C. PCR products were separated by 1.5% agarose gel electrophoresis and the DNA fragments were subcloned into the Bluescript SK+ plasmid. T7 and T3 primers were used to determine the double strand nucleotide sequences using LI-COR 4000 (Aloka Co., Ltd., Tokyo, Japan) or ABI PRISM™ 310 Genetic Analyzer (Applied Biosystems Japan Ltd., Tokyo, Japan).

RESULTS AND DISCUSSION

It has been suspected that the transgene(s) introduced into GM organisms are unstable and variable compared to the internal host genes. However, the neutral theory of molecular evolution[10] tells us that the rate of mutation, such as point mutations of nucleotide sequences, must occur equally in the whole genome of all organisms. As a result, elite organisms will be selected and prosper against environmental pressure. If the natural theory is applicable to GM crops, the rate of mutation of the transgene should be the same as that of internal host genes. However, the rate of substitutions and deletions of nucleotide and amino acid sequences of the transgene has not been investigated in detail. Here, the frequency of mutations of the transgenes and the internal host gene, *Cong* gene, was investigated in the individual genomes of 72 RR soybeans by de-

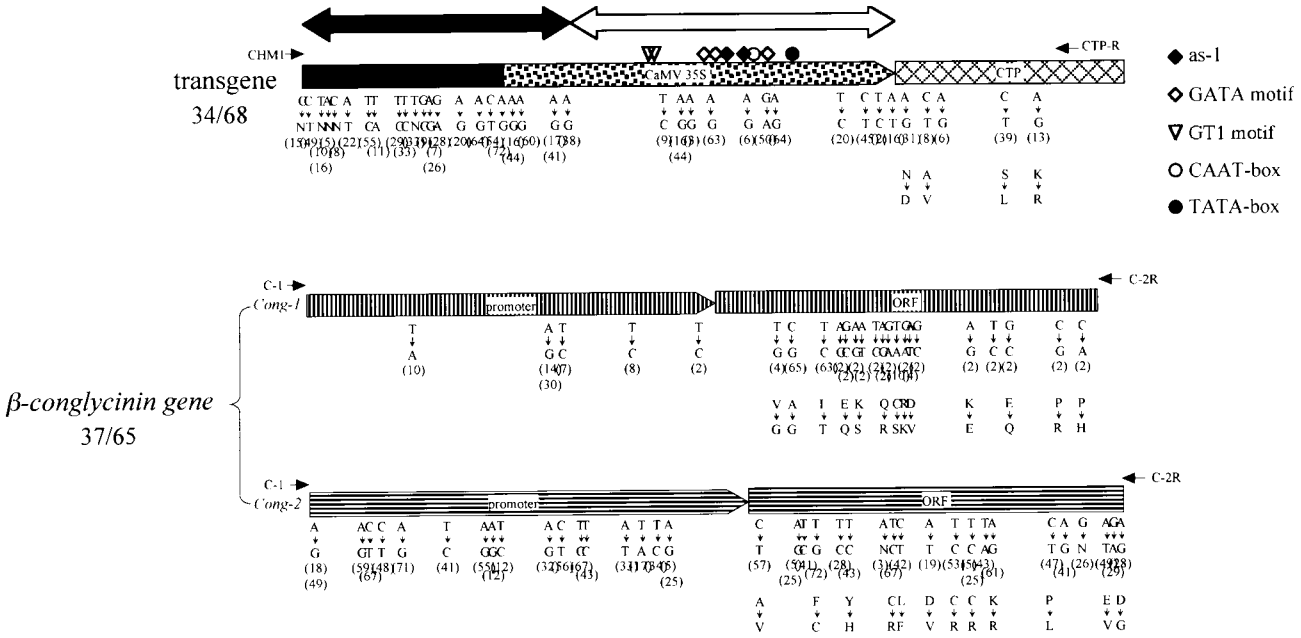

Fig. 1. Point Mutations and Deletions at the Proximal Promoter and Coding Regions of the Transgene and *Cong* Genes in Individual RR Soybeans

Among the 68 and 65 nucleotide sequences that could be determined by double strands, 34 and 37 differences could be observed in the transgene and *Cong* gene, respectively. The number in parentheses indicates the sample number of the individual RR soybeans in which substitutions or deletions of the nucleotide sequences were found. The arrowed upper letters, A, G, C and T, are nucleotides that were substituted to the lower letters. Nucleotides changed to N are point deletions. The lower parts under the coding regions (CTP and ORF) show the substitution of deduced amino acid sequences. The symbols at the upper side of the 35S promoter show the position of the *cis*-elements and CAAT- and TATA-boxes. The region with the open line and double arrowheads is the putative promoter region, while the closed line with double arrowheads indicates the upstream region of the promoter of the transgene.

termination of the nucleotide sequences of PCR amplified DNA fragments using the primer pairs, CHM1 and CTP-R for the transgene and C-1 and C-2R for *Cong* genes. Nucleotide sequences of 68, 15 and 50 of amplified DNA fragments for the transgene, *Cong-1* gene and *Cong-2* gene, respectively, were determined by double strand sequences (Fig. 1). The lengths of the determined sequences cut from the proximal region of the primer sites were 572 bp for the transgene, 604 bp for the *Cong-1* gene and 617 bp for the *Cong-2* gene (Table 1). Point mutations and deletions of 34 and 37 for the fragments derived from the transgene and *Cong* genes, respectively, were identified (Fig. 1). In total, we determined the nucleotide sequences of 38896 bps (68 fragments × 572 bps) and 39910 bps (15 fragments × 604 bps + 50 fragments × 617 bps) for the transgene and *Cong* genes, respectively. This showed that one mutation was found per 1144 bps in the transgene and 1079 bps in the *Cong* genes (Table 1). These results indicated that the mutation rate of the transgene was almost the same as that of the *Cong* genes. The amino acid sequences of the coding region of the transgene and *Cong* genes, 41 amino acids and 97 or 96 amino acids, respectively, were encoded in the fragments. Although only four amino acid substitutions were found in the transgene, 25 substitutions were found in the Cong proteins (Fig. 1). This showed that the substitution of amino acids was found per 697 and 250 amino acid sequences encoded in the transgene and *Cong* genes, respectively (Table 1). Despite almost the same chance to cause base substitutions between the transgene and *Cong* genes, amino acid substitutions were repressed in the transgene compared to those in the *Cong* genes. The repression of the substitution in the amino acid sequences was because the mutations in the nucleotide sequence of the coding region of the transgene occurred in the third letters of the codons. If the amino acid sequences of the transgene changed, some of the substitutions would result in a loss of the functions of the transit peptide, CTP. This would lead to a loss of the most important phenotype, glyphosate-tolerance. The loss of this phenotype would mean the loss of the superiority of RR soybeans to non-GM soybeans. Thus, during the development, maintenance and seed production of RR soybean lines, the highest selection pressure should be placed on preventing the loss of the glyphosate-tolerant phenotype. It is thought that the amino acid

Table 1. Frequency of the Nucleotide and Amino Acid Sequences of the Transgene and *Cong* Gene in Individual RR Soybeans

	Number of confirmed nucleotide sequences	Number of nucleotide sequences (bp)	Total number of nucleotide sequences (bp)	Total number of mutation points	Average length of appearance of the mutation (bp/one mutation)
Mutation on the nucleotide sequence					
Transgene	68	572	38896	34	1144
Cong-1 gene	15	604	39910	37	1079
Cong-2 gene	50	617			
	Number of confirmed nucleotide sequences	Number of deduced amino acid sequences	Total number of amino acid sequence	Total number of amino acid substitutions	Average length of the appearance of the mutation (amino acids/one substitution)
Mutation on the amino acid sequence					
Transgene	68	41	2788	4	697
Cong-1 gene	15	97	6255	25	250
Cong-2 gene	50	96			

sequence of the transgene might be tend to be more stable but less change than other genes.

An objection could be raised that the transgene is a new gene, which was introduced into the soybean genome about a decade ago.[3] However, *Cong* genes are old internal host genes that exist in the elite soybean inbred line(s) established several decades ago. Alterations of nucleotide and amino acid sequences have accumulated in *Cong* genes of the inbred lines, which were introduced into the original glyphosate-tolerant soybean by a backcross during breeding to develop the Roundup Ready® soybean. Cong proteins are storage proteins in soybean seeds and some of the alterations in the amino acid sequences do not cause lethal mutations. As such, alterations might have accumulated in the soybean genome (Fig. 1). The above results should be confirmed by future experiments to investigate the mutation rate in the transgene after several decades.

It is remarkable that the alteration of the nucleotide sequences of the promoter region of the *Cong-1* gene might be reduced in comparison with the coding region of the *Cong-1* gene and the promoter region of the *Cong-2* gene (Fig. 1). This reduction might prevent the loss of promoter activity in the *Cong-1* gene, which suggests that expression of the *Cong-1* gene might be required to produce Cong-1 protein as a storage protein in soybean seeds, even if several amino acid sequences of Cong-1 protein have been altered. Therefore, selection pressure by breeders might be placed on the promoter region, and not on the coding region, of the *Cong-1* gene, whereas it might be placed equally on both regions of the *Cong-2* gene. When the nucleotide sequence of the soybean genomic DNA and 35S promoter region was divided into the proximal promoter region (shown by the open line with double arrowheads in Fig. 1) and its upstream region (closed line with double arrowheads in Fig. 1), the number of mutation points in each region was found to be completely different (12 in about 255 bps and 25 in about 210 bps for the proximal and upstream regions, respectively). That is, the nucleotide sequence of the proximal promoter region was less mutated and more stable than the 5′ upstream region. The transgene in the original glyphosate-tolerant soybean should have a homologous nucleotide sequence. Divergence of the nucleotide sequences was detected in the progeny of the RR soybean. The 35S promoter is the most popular and common promoter to drive transgenes in transgenic plants because its expression is very high, but it is less tissue or organ specific in many plant species.[11] Several *cis*-elements important for its expression have been revealed.[12] These include as-1[13] and the GT-1[14] and GATA motifs.[15] All these elements were located in the proximal promoter region (see Fig. 1). If the point mutation occurred in the *cis*-element that included the essential boxes for transcription factors, *e.g.*, CAAT and TATA, it should stop or reduce promoter activity, which would result in a loss or reduction of the glyphosate-tolerant phenotype. Therefore, breeders should continually select highly tolerant line(s) during the reproduction of seeds to keep the nucleotide sequence of the important region of the 35S promoter. Although the region noticed here was a homologous nucleotide sequence to the original, the speed and rate of mutations were different between the proximal promoter and its upstream regions. Selective pressure (artificial inbreeding bias) from the developer and pro-

ducer of the seeds might lead to the nucleotide sequence of the important region for strong promoter activity being kept and not as greatly mutated than the upstream region of the promoter. This might ensure a high expression of the transgene, and perhaps result in phenotypes derived from it, glyphosate tolerance being kept and maintained as in the Roundup Ready® soybeans that are commercially superior to non-GM soybeans.

Acknowledgements This study was supported by a grant of MHLW of Japan.

Takeshi Ogasawara and Yukie Chikagawa contributed equally to this report as the first author.

Yoshio Itoh was deceased on July 2, 2004.

REFERENCES

1) Register, J. C., Peterson, D. J., Bell, P. J., Bullock, W. P., Evans, I. J., Frame, B., Greenland, A. J., Higgs, N. S., Jepson, I. and Jiao, S. (1994) Structure and function of selectable and non-selectable transgenes in maize after introduction by particle bombardment. *Plant Mol. Biol*., **25**, 951–961.
2) Kohli, A., Griffiths, S., Palacios, N., Twyman, R. M., Vain, P., Laurie, D. A. and Christou, P. (1999) Molecular characterization of transforming plasmid rearrangements in transgenic rice reveals a recombination hotspot in the CaMV 35S promoter and confirms the predominance of microhomology mediated recombination. *Plant J*., **17**, 591–601.
3) Padgette, S. R., Kolacz. K. H., Delannay, X., Re, D. B., LaVallee, B. J., Tinius, C. N., Rhodes, W. K., Otero, Y. I., Barry, G. F., Eichholts, D. A., Peschke, V. M., Nida, D. L., Taylor, N. B. and Kishore, G. M. (1995) Development, identification, and characterization of a glyphosate-tolerant soybean line. *Crop Sci*., **35**, 1451–1461.
4) Delannay, X., Bauman, T. T., Beighley, D. H., Buettner, M. J., Coble, H. D., DeFelice, M. S., Derting, C. W., Diedrick, T. J., Griffin, J. L., Hagood, E. S., Hancock, F. G., Hart, S. E., LaVallee, B. J., Loux, M. M., Lueschen, W. E., Matson, K. W., Moots, C. K., Murdock, E., Nickell, A. D., Owen, M. D. K., Paschall, E. H., Prochaska, L. M., Raymond, P. J., Reynolds, D. B., Rhodes, W. K., Roeth, F. W., Sprankle, P. L., Tarochione, L. J., Tinius, C. N., Walker, R. H., Wax, L. M., Weigelt, H. D. and Padgette, S. R. (1995) Yield evaluation of a glyphosate-tolerant soybean line after treatment with glyphosate. *Crop Sci*., **35**, 1461–1467.
5) Armstrong, C. L., Parker, G. B., Pershing, J. C., Brown, S. M., Sanders, P. R., Duncan, D. R., Stone, T., Dean D. A., DeBoer, D. L., Hart, J., Howe, A. R., Morrish, F. M., Pajeau, M. E., Petersen, W. L., Reich, B. J., Rodriguez, R., Santino, C. G., Sato, S. J., Schuler, W., Sims, S. R., Stehling, S., Tarochione, L. J. and Fromm, M. E. (1995) Field evaluation of European corn borer control in progeny of 173 transgenic corn events expressing and insecticidal protein from *Bacillus thuringiensis*. *Crop Sci*., **35**, 550–557.
6) Ogasawara, K. (2005) Development of detection method for genetically modified food. In *Ph D thesis of Tokyo University of Agriculture and Technology*, Tokyo, Japan.
7) Doyle, J. J., Schuler, M. A., Godette, W. D., Zenger, V., Beachy, R. N. and Slightom, J. L. (1986) The glycosylated seed storage proteins of *Glycine max* and *Phaseolus vulgaris*. Structural homologies of genes and proteins. *J. Biol. Chem*., **261**, 9228–9238.
8) Yoshino, M., Kanazawa, A., Tsutsumi, K., Nakamura, I. and Shimamoto, Y. (2001) Structure and characterization of the gene encoding alpha subunit of soybean β-conglycinin. *Genes Genet. Syst*., **76**, 99–105.
9) Yamaguchi, H., Sasaki, K., Kidachi, Y., Shirama, K., Kiyokawa, S., Ryoyama, K., Matsuoka, T., Hino, A., Umetsu, H. and Kamada, H. (2000) Detection of recombinant DNA in genetically modified soybeans and tofu. *Jpn. J. Food Chem*., **7**, 112–116.
10) Kimura, M. (1983) *The neutral theory of molecular evolution*. Cambridge University Press, London.
11) Bevan, M. W., Manson, S. E. and Goelet, P. (1985) Expression of tobacco mosaic virus coat protein by a cauliflower mosaic virus promoter in plants transformed by *Agrobacterium*. *EMBO J*., **4**, 1921–1926.
12) Odell, J. T., Nagy, F. and Chua, N.-H. (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature* (London), **313**, 810–812.
13) Lam, E., Benfey, P. N., Gilmartin, P. M., Fang, R. X. and Chua, N.-H. (1989) Site-specific mutations alter in vitro factor binding and change promoter expression pattern in transgenic plants. *Proc. Natl. Acad. Sci. U.S.A*., **86**, 7890–7894.
14) Lam, E. and Chua, N.-H. (1990) GT-1 binding site confers light responsive expression in transgenic tobacco. *Science*, **248**, 471–474.
15) Lam, E. and Chua, N.-H. (1989) ASF-2: A factor that binds to the cauliflower mosaic virus 35S promoter and a conserved GATA motif in *Cab* promoters. *Plant Cell*, **1**, 1147–1156.