

A tool for functional plant genomics: Chimeric RNA/DNA oligonucleotides cause *in vivo* gene-specific mutations

PETER R. BEETHAM, PETER B. KIPP, XENIA L. SAWYCKY, CHARLES J. ARNTZEN, AND GREGORY D. MAY*

Boyce Thompson Institute for Plant Research, Cornell University, Tower Road, Ithaca, NY 14853-1801

Contributed by Charles J. Arntzen, May 5, 1999

ABSTRACT Self-complementary chimeric oligonucleotides (COs) composed of DNA and modified RNA residues were evaluated as a means to (i) create stable, site-specific base substitutions in a nuclear gene and (ii) introduce a frameshift in a nuclear transgene in plant cells. To demonstrate the creation of allele-specific mutations in a member of a gene family, COs were designed to target the codon for Pro-196 of SuRA, a tobacco acetolactate synthase (ALS) gene. An amino acid substitution at Pro-196 of ALS confers a herbicide-resistance phenotype that can be used as a selectable marker in plant cells. COs were designed to contain a 25-nt homology domain comprised of a five-deoxyribonucleotide region (harboring a single base mismatch to the native ALS sequence) flanked by regions each composed of 10 ribonucleotides. After recovery of herbicide-resistant tobacco cells on selective medium, DNA sequence analyses identified base conversions in the ALS gene at the codon for Pro-196. To demonstrate a site-specific insertion of a single base into a targeted gene, COs were used to restore expression of an inactive green fluorescent protein transgene that had been designed to contain a single base deletion. Recovery of fluorescent cells confirmed the deletion correction. Our results demonstrate the application of a technology to modify individual genetic loci by catalyzing either a base substitution or a base addition to specific nuclear genes; this approach should have great utility in the area of plant functional genomics.

Genomics is currently a central component of plant biology research. The gene sequence of *Arabidopsis thaliana* will soon be available, and many other species are under study. Identification of new genes is occurring at a much faster pace than is the determination of their function. One missing technology in plant biology is the ability to selectively and reliably create site-specific “gene knockouts” or homologous recombination of genes of interest or of unknown function. The phenomena of gene silencing (e.g., antisense and cosuppression) provides a method for understanding gene function through the creation of transgenic crops expressing a gene sequence that silences the endogenous gene. However, this approach is not suitable for functional genomic studies of individual members of multigene families or genes with similar sequence, and is sometimes problematic in other technical aspects (1).

A technology currently being explored in prokaryotic and eukaryotic systems uses self-complementary chimeric oligonucleotides (COs) comprised of DNA and 2'-O-methyl RNA to target and mutate genes *in vivo*. These COs are designed to have one or more bases that do not pair with the endogenous gene sequence. This approach was successfully used to modify endogenous genes of mammalian cells (2–5) in a site-specific and genetically inheritable manner. Recently, Alexeev and Yoon (5) have demonstrated that permanent *in vivo* conversions result in phenotypic changes in mouse melanocytes. It has

been hypothesized that the mechanism by which COs induce mutation is via action of DNA repair enzymes that recognize mismatches between the targeted gene and the CO, which was designed to create a mismatch (6).

Sulfonylurea herbicides retard plant growth by inhibiting branched-chain amino acid biosynthesis by blocking acetolactate synthase (ALS) (7). ALS is encoded by a diallelic gene family in *Nicotiana tabacum* (an allotetraploid species). The herbicides are no longer toxic to plants that contain at least one mutated gene encoding an altered form of ALS. A mutation that causes an amino acid substitution at Pro-196 confers resistance to the herbicide chlorsulfuron (Glean, DuPont) but the enzyme remains sensitive to another herbicide, imazaquin (Scepter, American Cyanamid) (8–10).

In this manuscript, we aimed to determine whether the introduction of COs is sufficient to cause targeted mutations in a nuclear gene in plant cells. To answer this question, COs were designed to modify an ALS gene in tobacco and an inactivated transgene that was created by deletion of a single nucleotide of the gene encoding the green fluorescent protein (GFP).

MATERIALS AND METHODS

Maintenance of Tobacco Cell Suspension Cultures and Plant Transformation. Tobacco Nt-1 cell suspensions were maintained as shaker cultures (27°C, 200 rpm in a 250-ml flask) and transferred weekly to fresh suspension medium (CSM) containing Murashige and Skoog salts (GIBCO/BRL), 500 mg/liter Mes, 1 mg/liter thiamin, 100 mg/liter myo-inositol, 180 mg/liter KH₂PO₄, 2.21 mg/liter 2,4-dichlorophenoxyacetic acid (2,4-D), and 30 g/liter sucrose (pH 5.7). For solidified medium, 8 g/liter agar-agar (Sigma) was added before autoclaving. Cell suspensions maintained for biolistic delivery were subcultured weekly by transferring 1 ml of an established suspension culture into 49 ml of fresh liquid CSM.

Transgenic tobacco plants expressing a nontranslatable form of GFP under the control of the CaMV 35S promoter were generated by using a standard *Agrobacterium*-mediated plant transformation protocol (11).

CO Design, Synthesis, and Labeling. COs were designed based on the sequence of the tobacco ALS SuRA allele (GenBank accession no. X07644) and the sequence of GFP (12). CO ALS1 was designed with the Pro-196 codon CCA altered to contain a mismatch codon CAA (which encodes glutamine). A second CO (ALS2) was designed to substitute CTA (a leucine codon) at codon 196 of ALS SuRA (Fig. 1). COs were synthesized and purified according to ref. 3. The SuRA and SuRB alleles differ by only a single nucleotide in the region targeted by COs. A nonspecific CO (NSC1) was designed as a negative control by using the same general CO

Abbreviations: CO(s), chimeric oligonucleotide(s); ALS, acetolactate synthase; GFP, green fluorescent protein.

A Commentary on this article begins on page 8321.

*To whom reprint requests should be sent at the present address: The Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73402. e-mail: gdmay@noble.org.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

could only occur if the frameshift mutation in the transgene-encoding GFP was corrected.

RESULTS AND DISCUSSION

Introduction of Chimeric Oligonucleotides into Plant Cells.

COs were fluorescently labeled in initial experiments to assist in the optimization of delivery into plant cells by electroporation (Fig. 2) and microparticle bombardment. As detected by using confocal UV microscopy, rhodamine-tagged COs were observed within plant cells, and were preferentially localized in the nucleus. The details supplied in the experimental protocols report our optimized precipitation and biolistic parameters for CO introduction into plant cells.

COs ALS1 and ALS2 were introduced into Nt-1 cells by microparticle bombardment. The following controls were also included in our experimental design: nonbombarded cells; cells bombarded with gold microprojectiles from a precipitation in which COs were excluded; cells bombarded with an oligonucleotide (ALSD) of DNA sequence identical to CO ALS1 synthesized without RNA residues (i.e., an all DNA oligonucleotide); and cells bombarded with a nonspecific CO (NSC1).

After CO introduction, putative conversions to herbicide resistance were identified as cell masses actively growing on selective medium (Fig. 3). Table 1 summarizes the results of three separate experiments. It has been demonstrated that spontaneous mutations in ALS can be recovered in tobacco cells in culture (7) and that herbicide resistance may arise from independent, different amino acid substitutions in the enzyme (8–10). Amino acid substitutions at codons other than Pro-196 of the tobacco ALS gene can confer resistance to both sulfonylurea (Glean) and imidazolinolone (Scepter) herbicides, whereas a mutation at codon 196 leads to selective resistance to Glean but not Scepter (9, 10). To eliminate possible spontaneous mutations at sites in the gene other than Pro-196, candidate cell lines were grown on medium containing 3 ppm imazaquin (negative selection) (Table 1).

By measuring growth of the Glean-resistant cells on Scepter, it was possible to determine further the ratio of Glean-resistant/Scepter-sensitive cells recovered. Mutation to this phenotype was 20-fold higher in the samples bombarded with either COs ALS-1 or ALS-2 when compared with the untreated controls. This finding suggests that codon 196 had been selectively mutated. To obtain more precise evidence that COs ALS1 and ALS-2 were causing a selective mutation rate

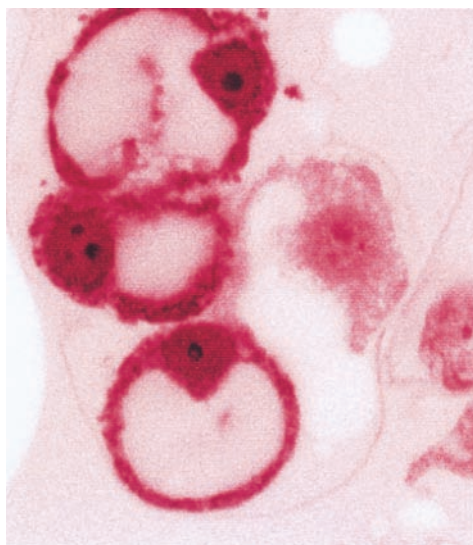


FIG. 2. Confocal photomicrograph demonstrating nuclear localization of rhodamine-tagged COs in tobacco Nt-1 cell protoplasts.



FIG. 3. Generation and selection of Glean-resistant Nt-1 cells grown on selective medium. Three-day-old plate cultures of Nt-1 cell suspension were used in ALS CO biolistic delivery experiments. Bombarded cell cultures were maintained on 15 ppb chlorsulfuron; putatively converted Nt-1 cell masses were transferred to medium containing 50 ppb chlorsulfuron.

increase at codon 196, PCR products that amplify this region of the ALS sequence were generated and sequenced.

It is important to note that there are four possible locations in the tobacco genome at which codon 196 can be mutated. That is, in the allotetraploid genome, Pro-196 could be mutated in either copy of SuRA or SuRB. Fig. 4 shows an electropherogram of the DNA sequence analyses of PCR products from Glean^R/Scepter^S cell lines. If an individual ALS gene at one locus had been mutated, but not the other ALS loci, it would be expected that a dual peak would be found in the PCR product at the point of mutation. This was observed at the codon from amino acid 196 (see arrow); we concluded that at least one gene of the ALS multigene family had been altered. To confirm this observation, PCR products were ligated into the plasmid pGEM-T(easy), amplified, and individually sequenced. Cloned PCR products from the control cell lines demonstrate wild-type sequence in every instance, whereas 12–67% of the cloned PCR products generated from Glean^R/Scepter^S cell lines contained a modified base at the codon for amino acid 196. The fact that the cloned PCR products derived from the SuRB allele did not contain a mutation can be interpreted as evidence that the COs ALS1 and ALS2 are acting selectively at the SuRA allele.

Interestingly, the modified base in the Pro-196 codon was always found to be one nucleotide 5' of the mismatch nucleotide of both COs. This heterotopic modification converted the Pro-196 CCA to a Thr-196 ACA (with CO ALS1) or Ser-196 TCA (CO ALS2) rather than the CAA or CTA as projected. The heterotopic modification of ALS by CO ALS2 had both ACA and TCA conversions. The phenomenon of a selective 5'-shift in mismatch "repair" (or, more correctly, the site of the observed mutation) may be indicative of the repair enzyme activities in plant cells under these experimental conditions. It has been previously reported that the formation of abasic sites may lead to nontargeted mutations in other systems (13, 14). For example, Bishop *et al.* (13) examined the miscoding properties of modified guanine residues bearing increasingly bulky O-6 substituents. Rat4 cells were transfected with plasmids carrying H-ras genes in which O-6-ethyl- and O-6-benzylguanine were substituted for the first or second

Table 1. Summary of ALS gene conversion experiment

Treatments	Plates, no.*	Glean ^R /Scepter ^S cell masses, no.†
Untreated	15	2 (0.1)
Microcarriers	15	1 (0.1)
CO NSC1	9	3 (0.3)
ALSD (DNA only)	9	2 (0.2)
CO ALS-1	9	21 (2.3)
CO ALS-2	9	19 (2.1)

*Number of plates used in experiment (One ml of packed cell volume bombardment).

†Number of Glean^R/Scepter^S cell masses obtained. Frequency per bombardment is shown in parentheses.

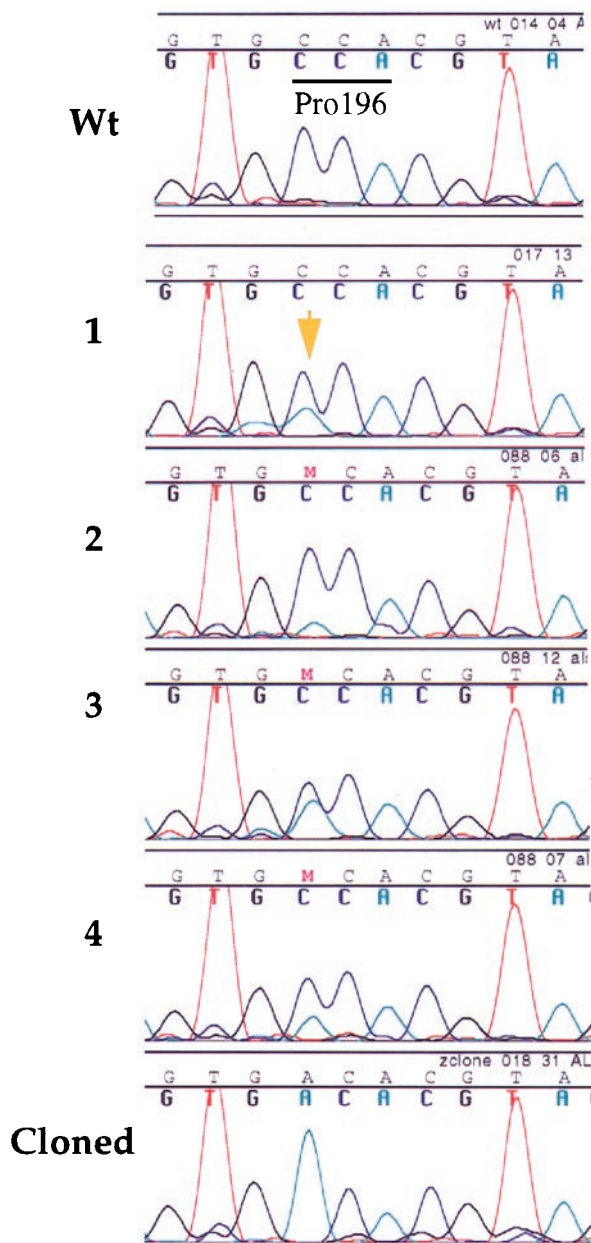


FIG. 4. Automated nucleotide sequence analyses of gene modification events. Electropherograms displaying the sequence of ALS-specific PCR products derived from DNA isolated from untreated Nt-1 cells (wt) and herbicide-resistant Nt-1 cells recovered as a result of an introduced CO (1-4) indicate an alteration in the native sequence for codon 196 (CCA). The conversion (C → A; arrow) in codon 196 (CCA) was detected (1-4) and is consistent with a targeted mutation induced by the CO. Because there are multiple copies of ALS alleles, modification of a single gene can confer herbicide resistance, but sequence analyses will indicate base composition heterogeneity at the revised position. To characterize individual species from the heterogeneous population of the ALS-specific PCR products derived from the DNA of herbicide resistant cells, shown in electropherograms 1-4, these products were cloned and sequenced. Sequence analysis of 1 of the 24 randomly selected clones, shown (Cloned), demonstrates the C → A conversion.

guanine residues of codon 12 (GGA). Their results demonstrate that either O-6-substituents induced "semitargeted" as well as targeted mutations. Semitargeted mutations were strictly G → A at the base 5' to a position 2 adduct. The authors postulated that the mechanism for nontargeted mutations may be related to abasic site formation or to translesion DNA

a

GFP1 CO **TACc**acucguucc**CG**CTCCucgacaagug
 GFP(Δ) **ATGGT**GAGCAAGGGC-**AGGAGCTGTT**CAC
 M V S K G **Frame-shift**
 Wt GFP **ATGGT**GAGCAAGGGC**GAGGAGCTGTT**CAC
 M V S K G E E L F T

b

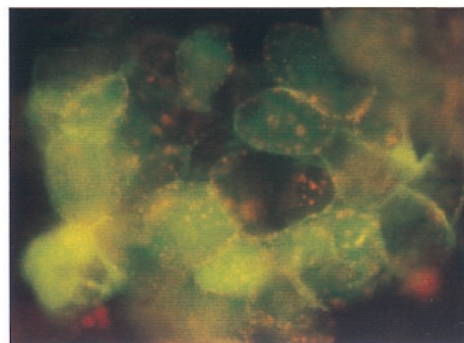


FIG. 5. Target sequence of the mutant GFP(Δ) and the GFP1 CO (a). Fluorescence photomicrograph of a putative conversion to the GFP phenotype in leaf callus from the tobacco GFP-mutant line Δ6 (b).

synthesis. In addition, Kamiya *et al.* (14) constructed c-Ha-ras genes with a true abasic site in codon 12. The ras genes were found to be activated in NIH 3T3 cells by a mutation at the modified site and, more frequently, at flanking positions.

Tobacco Nt-1 cells are allotetraploid with a genome of ≈4,000 megabase pairs (9). The ALS gene (SuRA allele) therefore is a target gene in a complex plant genome. CO bombardment resulted in herbicide-resistant cell lines that were shown to have nucleotide modifications at the codon encoding Pro-196. We have two lines of evidence demonstrating that the herbicide-resistant cells we recover do not arise from spontaneous mutations. First, in experiments using an equal number of Nt-1 cells, we have recovered 10- to 20-fold less herbicide-resistant cell lines after bombardment with uncoated gold microprojectiles, a nonspecific CO, or a "DNA-only" oligonucleotide. That is, the ALS COs increase mutation frequency by as much as a factor of 20. The fact that the increased mutation frequency is due to a site-specific mutation was verified by sequence analyses of ALS-specific PCR products (Fig. 4) amplified from the DNA of herbicide-resistant cells and indicates that these cells contain a base change specifically in the codon for Pro-196. PCR products subsequently cloned and sequenced confirm the presence of amplified DNAs with a modified Pro-196 codon. These modified cell lines have demonstrated genotypic stability by maintaining herbicide resistance over several months, and the modified Pro-196 codon can also be identified as a stable nucleotide change, by repeated PCR, cloning, and sequencing.

Our data of cell clumps/bombardment suggest the efficiency of base conversion at the ALS locus is up to two orders of magnitude higher than the controls, although the efficiency of CO-directed conversion was variable between experiments. These results are similar to the recent observations of CO-directed conversion in other systems (2, 3). Variation in efficiency may be caused by variation in CO delivery or cellular competency for conversion, which may vary with cell cycle or may be related to activity of DNA repair enzymes.

Further studies of other sequence targets may help to clarify the heterotopic nature of gene modification observed in this study. Interestingly, conversions directed by the ALS2 CO resulted in both C → T and C → A changes. Further characterization of plant DNA recombinase and repair enzymes may provide a better understanding of CO mode of action in plants.

Reactivation of an Inactive GFP Transgene. GFPs are a unique class of proteins involved in the bioluminescence of many jellyfish. In plants, GFP has been extensively used as a reporter for gene expression in both transient and stable expression systems (12). With the idea of using GFP as a model system to monitor conversion events in plant cells, we generated a mutant GFP expression vector that contained a single base pair deletion in the ORF that results in a frameshift mutation, thus preventing translation of the protein.

The CO GFP1 (designed to restore function to GFP) was introduced into the tobacco GFP-mutant line Δ6 by micro-particle bombardment, and recovery of the GFP phenotype was observed by using fluorescence microscopy (Fig. 5). Putative conversion events were identified on the basis of restoration of the GFP phenotype. No GFP expression was observed with the nonspecific CO controls. GFP expression was observed over a period of 3–10 days postbombardment.

This successful reactivation of the GFP gene provides evidence that COs can successfully catalyze the insertion of a nucleotide in a nuclear transgene that we had chosen for its visible marker phenotype. It suggests this “nucleotide insertion technology” would also be applicable for introduction of a base into actively expressed coding sequences, thus rendering the targeted gene inactive by causing a frameshift mutation.

Unlike antisense or cosuppression strategies, directed gene knockouts would be of value in inactivating specific members of plant multigene families; as long as a nonhomologous sequence in the members of the family can be identified for use in designing a CO. In addition to gene silencing, this technology could be used to alter untranslated or nontranscribed regions of a gene.

SUMMARY

The experiments presented herein have shown two different types of induced mutation that are catalyzed in a site-specific means by COs designed to hybridize to a unique 25-base sequence in the nuclear genome. First, the mutational frequency for alteration of a specific codon in a gene can be enhanced by as much as 20-fold. We anticipate that this mutational efficiency will be increased when the mechanism controlling the CO-induced mutation process is better characterized. Second, the coding sequence of a nuclear gene can be selectively altered by the insertion of an additional nucle-

otide; this is catalyzed specifically by the CO designed for the unique DNA sequence where the base was inserted, because the activation of the GFP was never observed in the absence of the appropriate CO.

A further experiment to combine the above-mentioned types of mutations can now be contemplated. If two different COs were both delivered into plant cells, it is hypothetically possible that a selectable mutation could be created in a plant cell together with a second mutation that would cause a frameshift inactivation of an unlinked gene. That is, dual mutations might be created in which one is known with respect to function, but the second could be created to explore the cellular consequence of loss of function of a unique coding sequence. If perfected, this approach could have great value in functional genomics. The approach could also have more immediate value in causing the inactivation or alteration of endogenous plant genes wherein the alteration could increase crop value.

We thank Ramesh Kumar, Eric Kmiec, Naomi Thomson, Keith Walker, Tong Zhu, Ben Bowen, and Chris Baszczynski (Pioneer Hi-Bred International, Inc.) for valuable discussions. This work was supported in part by the Boyce Thompson Institute for Plant Research, Inc. and also by a research contract from Kimeragen, Inc. (Newtown, PA).

1. Fink, G. R. (1998) *Genetics* **149**, 473–477.
2. Yoon, K., Cole-Strauss, A. & Kmiec, E. B. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 2071–2076.
3. Cole-Strauss, A., Yoon, K., Xiang, Y., Byrne, B. C., Rice, M. C., Cryn, J. & Kmiec, E. B. (1996) *Science* **273**, 1386–1389.
4. Kren, B. T., Bandyopadhyay, P. & Steer, C. J. (1998) *Nat. Med.* **4**, 285–290.
5. Alexeev, V. & Yoon, K. (1998) *Nat. Biotechnol.* **16**, 1343–1346.
6. Cole-Strauss, A., Gamper, H., Holloman, W. K., Muñoz, M., Cheng, N. & Kmiec, E. B. (1999) *Nucleic Acids Res.* **27**(5), 1323–1330.
7. Chalef, R. S. & Ray, T. B. (1984) *Science* **223**, 1148–1151.
8. Chalef, R. S. & Mauvais, C. J. (1984) *Science* **224**, 1443–1445.
9. Lee, K. Y., Townsend, J., Tepperman, J., Black, M., Chui, C. F., Mazur, B., Dunsmuir, O. & Bedbrook, J. (1988) *EMBO J.* **7**, 1241–1248.
10. Bedbrook, J. R., Chalef, R. S., Falco, S. C., Mazur, B. J., Somerville, C. R. & Yadav, N. S. (1991) U.S. Patent **5**, 013,659.
11. Horsch R. B., Fry J. B., Hoffmann N. L., Wallroth M., Eichholtz D., Rodgers S. B. & Fraley R. T. (1985) *Science* **227**, 1229–1231.
12. Chiu, W. L., Niwa, Y., Zeng, W., Hirano, T., Kobayashi, H. & Sheen, J. (1996) *Curr. Biol.* **6**(3), 325–330.
13. Bishop, R. E., Pauly, G. T. & Moschel, R. C. (1996) *Carcinogenesis* **17**, 849–856.
14. Kamiya, H., Suzuki, M. & Ohtsuka, E. (1993) *FEBS Lett.* **328**, 125–129.