

## Quantitative Analysis of Transcript Accumulation from Genes Duplicated by Polyploidy Using cDNA-SSCP

BioTechniques 34:726-734 (April 2003)

### ABSTRACT

Repeated rounds of polyploidy have been commonplace in the lineages leading to modern eukaryotic genomes, giving rise to widespread gene duplication. Genes duplicated by polyploidy, or "homoeologs," may continue to be expressed at equal levels following polyploidization or their expression may be dramatically altered. In this report, we describe how SSCP analysis of RT-PCR products can be used to evaluate the expression status (presence and relative quantity) of highly similar homoeologous gene pairs from an allotetraploid genome. This cDNA-SSCP approach was used to evaluate transcript abundance in "synthetic tetraploid" mRNA pools (i.e., mixtures of diploid products) and three natural homoeologous gene pairs expressed in tetraploid cotton (*Gossypium hirsutum*) ovules. Results from replicated tests show that cDNA-SSCP reliably separates duplicated transcripts with 99% sequence identity. Most significantly, the method yields quantitative estimates of transcript ratios in template pools that range from equimolar to approximately 100:1.

### INTRODUCTION

Most eukaryotic organisms have experienced one or more polyploidization events during their evolutionary history, which contributes to the genetic redundancy characteristic of many eukaryotic genomes (1–3). Polyploidy results in the simultaneous doubling of all genes, creating duplicated gene pairs, or homoeologs. The presence of two (or more) nonidentical genomes in a common nucleus creates a situation whereby transcript pools can be modified by altering the expression of one of the homoeologs. The resulting transcript pools may include an equal expression of both homoeologs or they

may show a bias for one of the two homoeologs (4,5).

To accurately assess the expression status of a single homoeologous gene pair, analytical methods must be able to resolve highly similar gene products [ $>98\%$  identical in polyploid cotton (6)] and to provide quantitative results on the relative abundance of each homoeolog. Several techniques have been successfully used to evaluate polyploid gene expression, but the methods most commonly used for transcript analysis—namely, RNA gel blot hybridizations and expressed sequence tag (EST)-based microarrays—frequently fail to discriminate between nearly identical homoeologous gene pairs. Amplification-based transcript scanning techniques such as cDNA-AFLP (7,8) can provide high-resolution transcript analysis with wide genome coverage. This technique has been used to gain remarkable insight into the global frequency of silencing in synthetic (9,10) and natural (11) polyploids; nevertheless, the method is not easily adapted to evaluate the expression status of a specific gene pair. Finally, the presence/absence of specific homoeologous pairs has been evaluated using RFLP analysis of RT-PCR products [RT-PCR/RFLP (12)]; however, this approach is only semi-quantitative (13), and it reveals nonidentical transcripts only when they differ by a restriction site.

Recently, Ibrahim et al. (14) used SSCP analysis to screen cDNA pools for the presence of multiple, non-homologous transcripts from the U1A spliceosomal protein gene family in the tetraploid potato. Here we show how SSCP approaches can be extended to characterize and quantitate the relative contribution of specific homoeologous gene pairs to the mRNA pool in a tetraploid plant, Upland cotton. This approach is sensitive and permits the discrimination of PCR products that differ by a single substitution (15–17). When coupled with phosphorimaging, cDNA-SSCP yields quantitative results of homoeologous gene expression across a range of relative transcript ratios from 1:1 (i.e., equal quantities of both homoeologs) to as great as 100:1.

### MATERIALS AND METHODS

#### Plant Materials and RNA Extractions

Upland cotton (*Gossypium hirsutum* L.;  $2N = 4X = 52$ ) contains two sets of homoeologous chromosomes, one that is similar to the modern A-genome diploid cotton *G. herbaceum* and a second that is similar to the D-genome species *G. raimondii* (18,19). Sequence divergence between A- and D-genome homoeologs of *G. hirsutum* and its diploid counterparts is minimal [0.4% to 0.6% per genome, respectively (6)], so *G. herbaceum* and *G. raimondii* were included in this study to optimize the separation of A- and D-genome amplification products and to assist in the identification of A- and D-genome homoeologs on SSCP gels.

Tetraploid *G. hirsutum* cv. TM1, diploid A-genome *G. herbaceum* (USDA accession A<sub>1</sub>-73), and diploid D-genome *G. raimondii* (Galau's) were grown in a greenhouse, and developing cotton ovules were harvested at 5, 10, 15, and 20 days post anthesis (dpa; flower opening). Ovular RNA was isolated using the method from Wan and Wilkins (20), and residual DNA was removed using DNA-free™ (Ambion, Austin, TX, USA). The synthesis of cDNA (Retroscrip™; Ambion) utilized RNA from either a single time point (10 dpa) or from a pool composed of equal quantities of 5–20 dpa RNA. These reactions included 6 µg total RNA in a 40-µL reaction volume or 2 µg RNA in a 20-µL volume.

#### RT-PCR of Homoeologous Loci

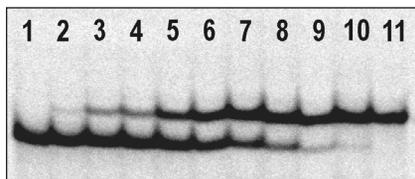
Following cDNA synthesis, PCR was used to amplify cDNAs of four low-copy genes, all of which have been sequenced from *G. hirsutum*, *G. herbaceum*, and *G. raimondii* (21). Locus "G3" is a putative 1,4-benzoquinone reductase (GenBank® accession nos. AY117105 and AY117106), and our PCR product included the terminal 305 bp of the gene when amplified with primers G3F (5'-GCATGTTTGAGATGGAGCAA-3') and G3R (5'-CTTGGCATTAAATCGA-AACCA-3'). A- and D-genome sequences of G3 from *G. hirsutum* are identical in size but differ by three substitutions,

which yields a pairwise nucleotide identity of 99%. The second gene, abbreviated “*F12*,” is a putative *cop1* interacting protein that yields a PCR product of 287 bp when amplified with primers F12F (5′-GAGACAACGGTCCATCCTGT-3′) and F12R (5′-GAAAACCCGA-AGCATAAGCA-3′). Homoeologous sequences of *F12* from *G. hirsutum* are identical in length and differ by six substitutions (97.9% identity; accession nos. AY116157-AY116158). The third gene, *myb1*, is a R2R3 MYB-domain transcription factor (1. when amplified with the primers cot105 (5′-AAGCAG-AGGAATTGATCCAC-3′) and cot106 (5′-ACCCTATGAATCCAAGGGTC-3′). Homoeologous sequences of *myb1* from *G. hirsutum* are identical in length and differ by 13 substitutions (97% identity; accession nos. AY115501 and AY115502). The final locus, alcohol dehydrogenase A (*adhA*) (23), was amplified using primers adhAF (5′-GGT-GAGGGTGTGAATGATTTG-3′) and adhAR (5′-GTACATATCCCACAGC-TAATGG-3′), which produced a 311-bp product. Homoeologous sequences of *adhA* from *G. hirsutum* are identical in length and differ by four substitutions (98.7% identity; accession nos. AF085064 and AF090146). PCR protocols followed previously described methods for cotton low-copy nuclear genes (6). Briefly, amplifications utilized 30 cycles of 1 min at 94°C, 1 min at 54°C–57°C, and 1 min at 72°C, with a final 72°C extension of 10 min. Reactions (30 µL) included 0.8 U *Taq* DNA polymerase and 1× buffer (Invitrogen, Carlsbad, CA, USA), 200 µM dNTPs, 1.5 mM MgCl<sub>2</sub>, 0.2 mM of each amplification primer, and 1 µL cDNA reaction. The amplification success was ver-

ified on 1.5% agarose gels, and the resulting PCR fragments were gel-purified for subsequent SSCP analysis (QIAquick™; Qiagen, Valencia, CA, USA).

### Homoeolog Labeling and SSCP Analysis

Following RT-PCR, 2–5 ng purified PCR product were used as a template for internal labeling. The amplification conditions were identical to those described earlier, except that (i) a single amplification primer was added to the reactions for all loci except *adhA*, which included both primers (see Results); (ii) 0.8 µCi [ $\alpha$ -<sup>32</sup>P]-dCTP (3000 Ci/mmol; Amersham Biosciences, Piscataway, NJ, USA) were added as an internal label; and (iii) the reaction volume was reduced to 10 µL. Following PCR cycling, 2 µL labeled product were mixed with 8 µL SSCP loading buffer (95% formamide, 20 mM EDTA, 10 mM NaOH, 0.05% bromphenol blue, 0.05% xylene cyanol), heated to 94°C for 3 min, and then cooled on ice. Four microliters of this mixture were loaded onto a 1× mutation detection enhancement gel (*MDE*; Cambrex Bio Science, Baltimore, MD, USA) containing 2% (*F12* and *myb1*) or 10% (*adhA* and *G3*) urea on a model S2 sequencer (Invitrogen). Running and gel buffers utilized 0.6× buffer containing Tris, Taurine, EDTA, pH 8.3. After the samples were loaded, gels were run at a 5 W constant for 16–18 h at an ambient temperature (*G3* and *F12*) or at 4°C (*adhA* and *myb1*). Following electrophoresis, the gels were dried on Whatman G3 paper and placed on phosphorimaging screens for 24–48 h. The images were visualized on a Storm™ 840 phosphorimager (Molecular Dynamics, Sunnyvale, CA, USA), and band intensities of homoeologous RT-PCR products were estimated using ImageQuant™ (version 5.2) software (Molecular Dynamics).



**Figure 1.** SSCP separation of A- and D-genome RT-PCR products of locus *G3* from synthetic tetraploid transcript pools. Lanes numbers correspond to the following transcript ratios: 1 = 100:0; 2 = 100:1; 3 = 20:1; 4 = 7:1; 5 = 3:1; 6 = 1:1; 7 = 1:3; 8 = 1:7; 9 = 1:20; 10 = 1:100; 11 = 0:100, for A- to D-genome RT-PCR products, respectively.

### Quantitative Analysis of Polyploid Transcript Pools

The quantitative nature of cDNA-SSCP was assessed through two experiments. RT-PCR products from locus *G3* were first amplified from *G. herbaceum* and *G. raimondii* 5–20 dpa

RNA pools, gel isolated, and quantitated by fluorometry. Synthetic A- and D-genome homoeolog mixtures were created by mixing A- and D-genome PCR products, respectively, in the following ratios (10  $\mu$ L volume): 4 ng A-genome (100:0); 4 ng A-genome + 0.04 ng D-genome (100:1); 4 ng A-genome + 0.2 ng D-genome (20:1); 4 ng A-genome + 0.57 ng D-genome (7:1); 4 ng A-genome + 1.33 ng D-genome (3:1); 2.5 ng A-genome + 2.5 ng D-genome (1:1); 1.33 ng A-genome + 4 ng D-genome (1:3); 0.57 ng A-genome + 4 ng D-genome (1:7); 0.2 ng A-genome + 4 ng D-genome (1:20); 0.04 ng A-genome + 4 ng D-genome (1:100); and 4 ng A-genome (0:100). Triplicate 3- $\mu$ L samples of these A- and D-genome mixtures were radiolabeled (as described earlier) and resolved by SSCP gel analysis. Natural transcript pools from *G. hirsutum* were then assayed by cDNA-SSCP for the presence of homoeologous transcripts from genes *F12*, *myb1*, and *adhA*.

The band intensities in both of these experiments were determined using either volumes (with background correction) or peak areas (10 measurements per lane). Tetraploid band volumes and peak areas were converted into a ratio of the proportion of A-genome to D-genome signals in the cDNA pool. Reported values represent an average of three independent labeling reactions (for the synthetic PCR pool) or three independent RT-PCRs and labeling reactions (for natural transcripts). The genomic identity of duplicate bands from *G. hirsutum* was assessed by co-migration with A- or D-genome diploid counterparts or by band isolation, followed by re-amplification and direct sequencing.

## RESULTS

### Quantitative Analysis of Synthetic and Natural Polyploid Transcript Pools

Figure 1 shows the results of one of three replicate SSCP gel analyses of A- and D-genome diploid cDNA mixtures for locus *G3*. SSCP resolution of diploid A-genome (lane 1) and D-genome (lane 11) RT-PCR products was sufficient to prevent band overlap,

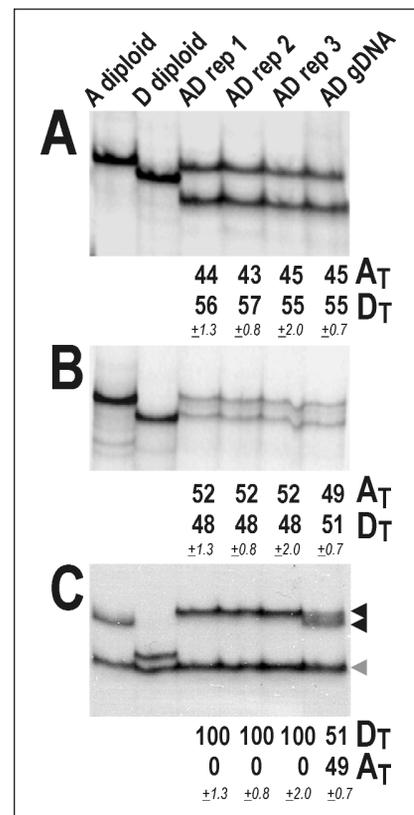
**Table 1. Quantitative Analysis of Relative Transcript Abundance for A- to D-Genome Diploid RT-PCR Products for Locus *G3***

Known Dilution Ratio	Fractional Ratio	Observed Signal Ratio ( $\pm$ sd)	95% Confidence Interval
0:100	0.000	-0.043 $\pm$ 0.042	-0.091–0.005
1:100	0.010	-0.011 $\pm$ 0.011	-0.025–0.002
1:20	0.048	0.046 $\pm$ 0.018	0.025–0.066
1:7	0.125	0.091 $\pm$ 0.017	0.071–0.110
1:3	0.250	0.299 $\pm$ 0.007	0.292–0.307
1:1	0.500	0.525 $\pm$ 0.016	0.507–0.543
3:1	0.750	0.740 $\pm$ 0.022	0.737–0.743
7:1	0.875	0.838 $\pm$ 0.030	0.796–0.879
20:1	0.952	0.918 $\pm$ 0.070	0.822–1.015
100:1	0.990	0.992 $\pm$ 0.026	0.956–1.029
100:0	1.000	1.035 $\pm$ 0.036	0.994–1.076

A- to D-genome RT-PCR products, respectively.

thereby simplifying the quantitative analyses of synthetic polyploid transcript mixtures (lanes 2–10). In general, observed A-genome:D-genome ratios showed close agreement with expected ratios, differing on average by 9.1% (Table 1; 0.21% to 27.4%). The 95% confidence intervals for observed A-genome:D-genome ratios include the known value in seven of 11 dilutions, while the remaining four dilutions (1:7, 1:3, 1:1, and 3:1; A-genome:D-genome, respectively) include the known value within the 90% confidence interval. While intermediate dilution ratios (e.g., from 1:7 to 7:1) appeared to show lower accuracy than other dilutions, they nevertheless showed the greatest precision, with unbiased coefficients of variation averaging 6.1% (2.2% to 20.2%). The high degree of accuracy and precision of the cDNA-SSCP method is reflected by the significant linear association between observed and known quantities (Spearman correlation coefficient,  $r_s = 0.998$ ;  $P < 0.001$ ; and slope = 1.019) in dilutions ranging from 100:1 to 1:100.

The broader utility of this method was evaluated by examining natural duplicate gene expression from *G. hirsutum* at loci *F12*, *myb1*, and *adhA* (Figure 2). Ovule-expressed *F12* showed SSCP migration patterns in which the A homoeolog migrated with its diploid relative, but the D homoeolog showed a higher mobility on SSCP than did diploid D-genome RT-PCR products (Figure 2A). In this example, A-



**Figure 2. Triplicate SSCP analysis of tetraploid gene expression in cotton ovules.** A and D indicate progenitor diploids, AD is the natural tetraploid *G. hirsutum*, and AD gDNA is a *G. hirsutum* genomic DNA control. Genes include locus *F12* (A), *myb1* (B), and *adhA* (C). Bolded numerical values show the percentage contribution of each genome ( $A_T$  or  $D_T$ ) to the total transcript pool; normal text shows the standard deviation of the estimate (expressed as a percent). Black arrows indicate SSCP bands primed by *adhAR*, while the gray arrow shows the band primed by *adhAF*.

# Short Technical Reports

---

genome homoeologs appeared less abundant than D-genome homoeologs (44:56), but this ratio was indistinguishable from the genomic DNA control (45:55). These results show that, while the amplification primers exhibit a subtle bias for D-genome templates, both genomes of the tetraploid contribute equally to the *F12* transcript pool. *Myb1* transcripts from *G. hirsutum* are also present in equal proportions to each other and to the genomic DNA control (Figure 2B). The notable lack of variation in the transcript ratios obtained for the replicate analysis of *F12* and *myb1* provides a measure of the high reproducibility of this method.

Initial SSCP results from *adhA* using a single amplification primer (*adhAF*) failed to produce discrete SSCP bands (data not shown). By repeating the labeling reaction with both primers included, a clearer polymorphism was detected on the alternatively

primed (*adhAR*) strand (Figure 2C). Genomic DNA controls showed that both genomic copies were amplified without apparent bias ( $A_T = 49\%$ ;  $D_T = 51\%$ ). However, in contrast to the results obtained from *F12* and *myb1*, the composition of homoeologous transcripts in developing ovules differed substantially from the genomic DNA control, revealing only a single band on the cDNA-SSCP. By directly sequencing tetraploid cDNA transcript pools from ovules and comparing patterns of expression from ovules to other tissues (e.g., roots and leaves; data not shown), we have determined that it is the A-genome copy of *adhA* that has been silenced in this organ.

## DISCUSSION

The cDNA-SSCP technique described here has several advantages

over cDNA-AFLP and RT-PCR/RFLP methods. For example, while cDNA-AFLP has the potential to sample hundreds of transcripts, the anonymous nature of the technique makes it difficult to adapt for evaluating the expression of pairs of homoeologous loci. In addition, nonidentical homoeologous transcripts that lack restriction site polymorphisms are invisible to both the cDNA-AFLP and RT-PCR/RFLP approaches. While cDNA-SSCP may be poorly suited for transcriptome-level expression studies of hundreds of genes, it may represent the ideal method for dissecting levels and patterns of known gene expression from polyploid organisms. The ability to resolve transcripts differing by a single nucleotide ensures that essentially any duplicate gene pair can be evaluated. Equal in significance, cDNA-SSCP produces a signal that is linear across a wide range of transcript ratios (Figure 1). The main

# Short Technical Reports

drawbacks of the SSCP-based approach are that gel conditions frequently require optimization for adequate band separation (17), and both strands may need to be screened to reveal a detectable polymorphism (see Figure 2C).

While cDNA-SSCP shows great promise [we have screened 40 homoeologous gene pairs in polyploid cotton using this method (13,24)], our study makes assumptions that are valid for tetraploid cotton but may not be warranted for all genes or organisms. First, we assumed that the ratio of RT-PCR products derived from homoeologous cDNAs would be equal to their ratio in the transcript pool. This assumption is likely to be robust when amplification targets and primers have a perfect match and when PCR is conducted within the linear range of amplification. Although our amplifications were conducted outside of the linear range of PCR, the sampling of homoeologous targets appears quantitative because the analysis of RT-PCR products from A- and D-genome diploids yielded observed ratios nearly identical to known values (Table 1). Also, SSCP bands derived from genomic DNA were present in a 1:1 ratio for *myb1* and *adhA* and a nearly 1:1 ratio for locus *F12* (Figure 2); these values match genomic copy numbers determined by Southern hybridization (data not shown). Together, these controls provide critical insight in revealing potential PCR bias that arises from primer mismatches, sequence-specific attributes (such as regions of self-complementarity or RNA secondary structure), or phenomena arising from amplification in the nonlinear range of PCR. Because inferences of transcript ratios may be significantly influenced by these phenomena, we suggest that titration experiments (Figure 1) and genomic DNA controls should be included in all cDNA-SSCP experiments.

Secondly, we assumed that tetraploid RT-PCR products would frequently co-migrate with RT-PCR products derived from A- and D-genome diploid cottons, thus facilitating homoeolog identification. While this was true for *G3* (data not shown) and *myb1*; it was not the case for *F12* and *adhA* (Figure 2, panels B and C). Given the unpredictable nature of SSCP band migration (15,17), backup methods

should be anticipated to verify the identity of loci. For example, to confirm that the *adhA* transcript pool was composed of D-genome PCR products (rather than co-migrating A- and D-genome products), we sequenced this product to confirm that it lacked additive (A+D) polymorphism. Such strategies will be a necessity when closely related diploids are unavailable or have yet to be identified.

Finally, we assumed that all loci would be homozygous within *G. hirsutum*. This assumption is reasonable for Upland cotton [given its low genetic diversity and high rate of self-fertilization (19)], but it cannot be expected for all polyploid species. Heterozygosity at a locus (or higher ploidy) will yield additional bands on SSCP, further obscuring locus identities. However, even in the more challenging cases, cDNA-SSCP may provide sufficient resolution to screen transcript pools for sequence homogeneity and sequence identity (14). Finally, the quantitative nature of this screening method can be used to rapidly identify unique variants that show differences in transcript abundance, regardless of ploidy level.

## REFERENCES

1. Ohno, S. 1970. Evolution by Gene Duplication. Springer-Verlag, New York.
2. Soltis, D.E. and P.S. Soltis. 1993. Molecular data and the dynamic nature of polyploidy. *Crit. Rev. Plant Sci.* 12:243-273.
3. Wendel, J.F. 2000. Genome evolution in polyploids. *Plant Mol. Biol.* 42:225-249.
4. Chen, Z.J., L. Comai, and C.S. Pikaard. 1998. Gene dosage and stochastic effects determine the severity and direction of uniparental ribosomal RNA silencing (nucleolar dominance) in *Arabidopsis* allopolyploids. *Proc. Natl. Acad. Sci. USA* 95:14891-14896.
5. Pikaard, C.S. 2001. Genomic change and gene silencing in polyploids. *Trends Genet.* 17:675-677.
6. Cronn, R.C., R.L. Small, and J.F. Wendel. 1999. Duplicated genes evolve independently after polyploid formation in cotton. *Proc. Natl. Acad. Sci. USA* 96:14406-14411.
7. Bachem, C.W.B., R.J.F.J. Oomen, and R.G.F. Visser. 1998. Transcript imaging with cDNA-AFLP: a step-by-step protocol. *Plant Mol. Biol. Rep.* 2:157-173.
8. Bachem, C.W.B., R.S. van der Hoeven, S.M. de Bruijn, D. Vreugdenhil, M. Zabeau, and R.G.F. Visser. 1996. Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. *Plant J.* 9:745-753.
9. Comai, L., A.P. Tyagi, K. Winter, R. Holmes-Davis, S. Reynolds, Y. Stevens, and B. Byers. 2000. Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. *Plant Cell* 12:1551-1568.
10. Kashkush, K., M. Feldman, and A.A. Levy. 2002. Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. *Genetics* 160:1651-1659.
11. Lee, H.-S. and Z.J. Chen. 2001. Protein-coding genes are epigenetically regulated in *Arabidopsis* polyploids. *Proc. Natl. Acad. Sci. USA* 98:6753-6758.
12. Song, K. and T.C. Osborn. 1994. A method for examining expression of homologous genes in plant polyploids. *Plant Mol. Biol.* 26:1065-1071.
13. Cedroni, M.E., R.C. Cronn, K.L. Adams, T.A. Wilkins, and J. F. Wendel. 2003. Evolution and expression of MYB genes in diploid and polyploid cotton. *Plant Mol. Biol.* 51:313-325.
14. Ibrahim, A.F.M., J.A. Watters, and J.W.S. Brown. 2001. Differential expression of potato U1A spliceosomal protein genes: a rapid method for expression profiling of multigene families. *Plant Mol. Biol.* 45:449-460.
15. Orita, M., H. Iwahara, K. Hayashi, and T. Sekiya. 1989. Detection of polymorphisms of DNA by gel electrophoresis as single-strand conformation polymorphism (SSCP). *Proc. Natl. Acad. Sci. USA* 91:2766-2770.
16. Sheffield, V.C., J.S. Beck, A.E. Kwitek, D.W. Sandstrom, and E.M. Stone. 1993. The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. *Genomics* 16:325-332.
17. Yip, S.P., D.A. Hopkinson, and D.B. Whitehouse. 1999. Improvement of SSCP analysis by use of denaturants. *BioTechniques* 27:20-24.
18. Brubaker, C.L., A.H. Paterson, and J.F. Wendel. 1999. Comparative genetic mapping of allotetraploid cotton and its diploid progenitors. *Genome* 42:184-203.
19. Wendel J.F. and R.C. Cronn. 2003. Polyploidy and the evolutionary history of cotton. *Advances in Agronomy* 78:139-186.
20. Wan, C.-Y. and T.A. Wilkins. 1994. A modified hot borate method significantly enhances the yield of high-quality RNA from cotton (*Gossypium hirsutum* L.). *Anal. Biochem.* 223:7-12.
21. Senchina, D.S., I. Alvarez, R.C. Cronn, B. Liu, A.H. Paterson, J. Rong, R.D. Noyes, R.A. Wing, et al. Rate variation among nuclear genes and the age of polyploidy in *Gossypium*. *Mol. Biol. Evol.* (In Press.)
22. Loguercio, L.L., J.Q. Zhang, and T.A. Wilkins. 1999. Differential regulation of six novel MYB-domain genes defines two distinct expression patterns in allotetraploid cotton (*Gossypium hirsutum* L.). *Mol. Gen. Genet.* 261:660-671.
23. Small, R.L. and J.F. Wendel. 2000. Copy number lability and evolutionary dynamics of the *Adh* gene family in diploid and tetraploid cotton (*Gossypium*). *Genetics* 155:1913-1926.
24. Adams, K., R. Cronn, R. Percifield, and J.F.

# Short Technical Reports

---

**Wendel.** Genes duplicated by polyploidy show unequal contributions to the transcriptome and silencing that are reciprocal and organ-specific. *Proc. Natl. Acad. Sci. USA* (In Press.)

*The authors contributed equally to the application of this method and manuscript preparation. We thank Ines Alvarez, Bao Liu, Ryan Percifield, and David Senchina for assistance in evaluating the loci selected, and especially Jonathan Wendel for providing a stimulating laboratory environment to develop this technique. This work was supported by a U.S. Department of Agriculture (USDA) National Research Initiative Grant and the Iowa State University Plant Sciences Institute (to J. Wendel) and a grant from the Jane Coffin Childs Memorial Fund for Medical Research (to K.A.). Trade names and commercial products are mentioned solely for information and no endorsement by the USDA is implied. Address correspondence to Dr. Richard Cronn, Pacific Northwest Research Station, 3200 SW Jefferson Way, Corvallis, OR, 97331, USA. e-mail: rcronn@fs.fed.us.*

Received 31 October 2002; accepted 15 January 2003.

**Richard C. Cronn and Keith L. Adams<sup>1</sup>**

*Pacific Northwest Research Station*

*U.S. Forest Service  
Corvallis, OR*

*<sup>1</sup>Iowa State University  
Ames, IA, USA*

## **RNA Interference by Production of Short Hairpin dsRNA in ES Cells, Their Differentiated Derivatives, and in Somatic Cell Lines**

*BioTechniques* 34:734-744 (April 2003)

### **ABSTRACT**

*dsRNA of several hundred nucleotides in length is effective at interfering with gene expression in mouse oocytes, pre-implantation embryos, and embryonic stem (ES) cells but is not as efficient in differentiated cell lines. Here we describe a method to achieve RNA interference in totipotent and differentiated ES cells together with a wide range of other mammalian cell types that is both simple and efficient. It utilizes a linearized plasmid that directs the expression of a hairpin RNA with a 22-nucleotide-paired region. This molecule has a 13-nucleotide 5' overhang that would be subject to capping on its 5' phosphoryl group and thus differs from the ideal structure suggested for effective small interfering RNAs. Thus, it appears either that the structure of small inhibitory RNA molecules may not need to be as precise as previously thought or that such a transcript is efficiently processed to a form that is effective in interfering with gene expression.*

### **INTRODUCTION**

The specific blocking gene expression by double-stranded RNA (dsRNA) of up to several hundred nucleotides in length has been documented in a variety of organisms including *Caenorhabditis elegans* (1), *Drosophila melanogaster* (2), trypanosomes (3), planaria (4), *paramecium* (5), hydra (6), zebrafish (7,8), and mouse (9,10). The use of dsRNA of such molecular weight for silencing gene expression has not been as efficient in cultured mammalian somatic cell lines where it appears to also inhibit gene expression nonspecifically (11). This reflects the ability of dsRNA to trigger the interference response in mammalian cells, lead-