The *Journal of Cotton Science* 4:65-69 (2000)
http://journal.cotton.org, © The Cotton Foundation 2000

BREEDING & GENETICS

Linkage Analysis of Transgenes Inserted into Cotton via *Agrobacterium tumefaciens* Transformation

Russell J. Kohel,* Jerry E. Quisenberry, Greg Cartwright, and John Yu

INTERPRETIVE SUMMARY

Linkage analysis was conducted between 2,4-D transgenes and 14 marker genes. Linkage was found between 2,4-D resistance and the naked seed-1 morphological marker (20–30 cM). Only two-point linkage tests were possible, so the orientation on the chromosome with respect to the marker could not be determined. Tests with aneuploid stocks deficient for the short arm of chromosome 12 placed the gene in the long arm of chromosome 12.

Knowledge of the location of the inserted gene would provide the opportunity to develop germplasm and strategies to maximize the efficiency of developing improved germplasm. Transformation is considered to be a random event, with each inserted gene at a unique location. So, each transgenic germline is the result of a separate insertion event, and the location of the insertion in the genome could represent different degrees of success in attempts at backcross improvement, depending on associated linkages.

ABSTRACT

The location of transgenes inserted into a genome are important in genetic studies and breeding programs. We conducted linkage analysis between 2,4-D resistant transgenes and 14 morphological marker genes in upland cotton (*Gossypium hirsutum* L.). Two separate germlines that exhibited monogenic dominance for resistance to 2,4-D were selected for linkage analysis. Multiple marker lines T582 and T586 were crossed with the 2,4-D resistant lines. Their F_1, F_2, and backcross/testcross progeny were produced and evaluated for segregation of resistance to 2,4-D and the marker loci. Linkage was found between 2,4-D resistance and the naked seed-1 morphological marker (18–37 cM), for both lines. Only two-point linkage tests were possible, so the orientation on the chromosome with respect to the marker could not be determined. Tests with aneuploid stocks deficient for the short arm of chromosome 12 placed the gene in the long arm of chromosome 12.

Transformation of cotton by *Agrobacterium tumefaciens*-mediated transformation is a documented technology in cotton (Firoozabady et al., 1987; Umbeck et al., 1987), as it is in other dicots (Assaad and Signer, 1992; Matzke et al., 1993; Misra, 1990; Misra and Gedamu, 1989; Puonti-Kaerlas et al., 1992; Scheid et al., 1991). Transgenes can be inserted as single events, are stable, and segregate with Mendelian expectations. Multiple insertion events occur, and they can result in unexpected interactions of gene expression (Scheid et al., 1991). Even unusual interactions can occur between single insertion events at separate sites when recombined sexually (Matzke et al., 1993). The typical procedure in improved cultivar development is to select and utilize single-event gene insertions.

Knowledge of the location of the inserted gene would provide the opportunity to develop germplasm and strategies to maximize the efficiency of developing improved germplasm. Transformation is considered to be a random event, with each inserted gene at a unique location. So, each transgenic germline is the result of a separate insertion event, and the location of the insertion in the genome could affect future success in attempts at backcross improvement, depending on associated linkages.

R.J. Kohel and J. Yu, USDA-ARS-SCRL, Crop Germplasm Res. Unit, 2765 F&B Road, College Station, TX 77845; J.E. Quisenberry, U.S. Pacific Basin Agric. Res. Center, P.O. Box 4459, Hilo, HI 96720, G. Cartwright USDA-ARS-CSRL, Plant Stress and Germplasm Dev. Res. Unit, Route 3, Box 215, Lubbock, TX 79401. Received 28 Oct. 1999. *Corresponding author (rjk0339@acs.tamu.edu).

Abbreviations: 2,4-D, 2,4-dichlorophenoxyacetic acid; T-DNA, transfer DNA.

MATERIALS AND METHODS

Our study of linkage relations between a gene inserted through transformation and classical morphological markers began with genetic material derived from studies by Bayley et al. (1992). The cotton in that study was transformed with a bacterial gene [2,4-D monooxygenase (*tfd*A)] to provide resistance to 2,4-D. Then those scientists inserted the gene by *A. tumefaciens*-mediated transformation. The resulting tissue cultures were selected for transformed cells, and calli of the transformed tissue were grown, induced to regenerate, and produced plantlets from each germline. Each transformed germline originated from callus produced on separate hypocotyl sections. Plantlets that were verified to contain the T-DNA for resistance to 2,4-D were grown, testcrossed to susceptible cottons, and self-pollinated. Seedlings from the testcrosses and self-pollinated seed were screened for resistance to 2,4-D and segregation of the transgene inserts.

Bayley et al. (1992) used hypocotyl tissue of 'Coker 312' for transformation because it is one of the few germplasms that has the potential for regeneration of plantlets from tissue culture. Coker 312 does not possess the desirable agronomic traits of contemporary cultivars, so transgenic regenerates must be backcrossed to desirable cultivars for use as an improved germplasm.

The gene for 2,4-D monooxygenase was introduced into cotton (Bayley et al., 1992), and we will refer to these transgenes as Ot_n in this paper. Regenerants that contained the target T-DNA were grown, crossed to an experimental line, RQSX-1, and F_1 seed were obtained. The progeny were screened for 2,4-D resistance segregation, and regenerants that segregated as monogenic for Ot_n were retained for further study.

The 2,4-D resistance trait is expressed as a complete dominant with no apparent difference in expression between one or two doses of the gene (Bayley et al., 1992). Six lines with single copy insertions were included in these experiments. Four lines were used as resistant controls and two were included in linkage tests (with Ot_4 and Ot_6 designating putatively separate 2,4-D resistant loci) to determine whether they were linked to any of the morphological markers; and if they were, with which linkage groups or chromosomes they were associated. All six lines with the gene insertions were recovered from germlines that originated from separate hypocotyl segments; therefore, they represented separate insertion events. It was assumed that they would segregate independently.

For the analysis of genetic linkage relations, transgenic lines were crossed with T582, multiple recessive marker line, and T586, multiple dominant marker line.

The T582 line includes the recessive marker loci-virescent-1, v_1 , L.G. XVII on chromosome 20; cup leaf, *cu* , location unknown; glandless-1, gl_1 , location unknown; frego bract, *fg* , L.G. VI on chromosome 3; and cluster-1, cl_1 , L.G. III on chromosome 16.

The T586 line includes the dominant marker loci-red plant, R_1 , L.G. III on chromosome 16; okra leaf, $L_2{}^{\circ}$, L.G. II on chromosome 15; tomentum, T_1 , L.G. IV on chromosome 6; petal spot, R_2 , L.G. I on chromosome 7; yellow pollen, P_1 , L.G. XI on chromosome 5; yellow petals, Y_1 , L.G. XII on an unknown A chromosome; brown lint, Lc_1 , L.G. I on chromosome 7; green lint, *Lg* , L.G. II on chromosome 15; and naked seed, N_1 , L.G. V on chromosome 12 (Endrizzi et al., 1984).

Segregation and linkage data analyses were conducted with the computer program G-MENDEL (Liu and Knapp, 1991).

Six 2,4-D resistant lines were grown and evaluated, and the two designated Ot_4 and Ot_6 were used in the crosses with the multiple marker lines. A normal tester, RQSX-1, and the two multiple marker lines, T582 (recessive) and T586 (dominant) were grown as the normal 2,4-D susceptible lines. Reciprocal F_1 crosses were grown for each multiple marker line: [(Ot_4 × T582); (T582 × Ot_4); (Ot_6 × T582); (T582 × Ot_6); (Ot_4 × T586); (T586 × Ot_4); (Ot_6 × T586); (T586 × Ot_6)]; and F_2s of each F_1 were grown. Reciprocal testcrosses of Ot_n × T586 and RQSX-1: [(RQSX-1(T586 × Ot_4)); ((T586 × Ot_4)RQSX-1); (RQSX-1(Ot_6 × T586)); ((T586 × Ot_6)RQSX-1)], and reciprocal backcrosses of Ot_n × T582 and T582: [(T582(T582 × Ot_4)); ((T582 × Ot_4)T582); (T582(Ot_6 × T582)); ((T582 × Ot_6)T582)] were grown. In some combinations with T582 no gl_1 segregated. The T582 line used in the original crosses was uniformly homozygous for all loci, except that it was segregating at the gl_1 locus and had likely mutated.

Fig. 1. Resistant transgenic cotton seedling, left, and susceptible control 3 d after a 0.3 g L^{-1} spray treatment of 2,4-D.

Seeds of each population were germinated in the greenhouse in peat pellets following the normal procedures of the genetics program at College Station, TX. Once the seedlings grew to the point where the first leaf was expanded, the seedlings were treated with 2,4-D (Fig. 1). A solution of 0.3 g L^{-1} was sprayed on the seedlings with an atomizer. In 2 weeks the seedlings were scored for morphological mutants and reaction to 2,4-D.

At 3 weeks of age the greenhouse grown seedlings were transplanted to field plots. Plots consisted of rows with 0.5 m spacing between plants within rows and 1.0 m spacing between rows. Field plants were verified for reaction to 2,4-D treatment and the genetic markers.

Tests were conducted in the greenhouse on those materials segregating for only the Ot_n genes. The seedlings were treated in the same manner as those seedlings that were to be transplanted to the field, except that after scoring for the reaction to 2,4-D the plants were treated again to ensure that there were no escapes. Crosses were made between Ot_4 and Ot_6, and the F_2 progeny were treated with 2,4-D and classified for resistant or susceptible seedling response. To test for chromosome location, aneuploid plants of known cytotype were crossed as female with the transgenic lines Ot_4 and Ot_6. The F_1 of these lines were classified for aneuploid vs. normal phenotype. These plants were self-pollinated, and the F_2 were treated and classified for reaction to 2,4-D. F_2 progeny from disomic F_1 plants or those independent of the aneuploid segment segregated 3 resistant to 1 susceptible. Progeny from aneuploid plants in which the inserted segment corresponded to the region missing in the aneuploid would be all resistant with no segregation.

Table 1. Segregation and analysis of F_2 and backcross/testcross (BC/TC) families from 2,4-D resistant transgenics (Ot_n) † crossed with susceptible cottons.

Family	Segregation (n)			
	Resistant	Susceptible	X^2	$P > G$ ‡
F_2^4			(3:1)	
Ot_4 × T582	44	16	0.09	0.77
T582 × Ot_4	46	12	0.57	0.44
Ot_4 × T586	44	17	0.27	0.61
T586 × Ot_4	46	10	1.52	0.20
Pooled	180	55	0.32	0.57
Heterogeneity (df = 3)			2.14	0.52
F_2^6				
Ot_6 × T582	39	16	0.47	0.49
T582 × Ot_6	46	9	2.40	0.12
Ot_6 × T586	48	15	0.05	0.83
T586 × Ot_6	48	21	1.04	0.31
Pooled	181	61	0.01	0.94
Heterogeneity (df = 3)			3.96	0.27
BC^4/TC^4			(1:1)	
T582 (T582 × Ot_4)	29	28	0.02	0.90
(T582 × Ot_4) T582	35	33	0.06	0.81
RQSX-1 (T586 × Ot_4)	31	16	4.79	0.03
(T586 × Ot_4) RQSX-1	34	36	0.06	0.81
Pooled	129	113	1.06	0.30
Heterogeneity (df = 3)			3.86	0.27
BC^6/TC^6				
T582 (Ot_6 × T582)	30	25	0.46	0.50
(T582 × Ot_6) T582	42	22	6.16	0.01
RQSX-1 (Ot_6 × T586)	33	34	0.01	0.90
(T586 × Ot_6) RQSX-1	19	24	0.58	0.44
Pooled	124	105	1.56	0.21
Heterogeneity (df = 3)			5.65	0.13

† *Ot* allele.

‡ G = approximation to Chi Square, $P > G$ = probability of a greater G value.

RESULTS AND DISCUSSION

Young seedlings in the greenhouse that expressed the cup leaf and frego bract traits could not be reliably scored for the reaction to 2,4-D treatment. The seedlings were all transplanted to the field and the 2,4-D reaction recorded as the plants grew. The reaction to 2,4-D ranged from plants completely misshapen to those with only a few leaves showing the symptoms. All the plants outgrew the effects of 2,4-D except for some stunting in the most severely damaged plants. Segregation of the F_2 and backcross families of the two transgenic lines conformed to the expectations of 2,4-D resistance being inherited as a single completely dominant gene (Table 1). There were no significant deviations in the individual families, the pooled data, or heterogeneity analyses in the F_2 and backcross populations of both transgenic loci.

Table 2. Linkage tests of transgenic loci (Ot_4 and Ot_6) in F_2 and backcross/testcross populations from crosses with T582 and T586 multiple marker cotton lines. †

Marker locus	F_2			Backcross/testcross		
	RC% ± SE ‡	G §	P > G	RC% ± SE	G	P > G
			Ot_4 †			
v_1	22.8 ± 8.2	0.03	0.87	50.5 ± 0.9	0.01	0.93
cu	40.7 ± 5.2	3.02	0.22	57.0 ± 3.4	2.03	0.15
gl_1	21.1 ± 8.3	1.99	0.16	44.2 ± 3.1	0.91	0.34
fg	26.8 ± 8.1	3.51	0.06	57.9 ± 3.6	2.55	0.11
cl_1	32.1 ± 7.8	3.88	0.05	68.2 ± 5.5	15.00	0.00
R_1	47.0 ± 5.6	4.92	0.09	52.8 ± 2.1	0.39	0.53
$L_2{}^o$	48.1 ± 5.6	8.00	0.02	50.4 ± 0.8	0.01	0.93
T_1	36.9 ± 7.8	1.46	0.23	56.8 ± 3.3	2.25	0.13
R_2	50.3 ± 6.9	0.61	0.43	50.4 ± 0.8	0.00	0.98
P_1	48.5 ± 7.0	0.01	0.94	43.9 ± 3.1	1.81	0.18
Y_1	52.1 ± 6.8	0.43	0.51	48.0 ± 1.2	0.21	0.65
Lc_1	48.3 ± 7.1	1.31	0.25	54.1 ± 2.6	0.84	0.36
Lg	43.9 ± 7.5	0.70	0.40	50.8 ± 1.1	0.01	0.92
N_1	30.4 ± 8.3	26.4	0.00	18.0 ± 7.2	57.96	0.00
			Ot_6			
v_1	23.9 ± 8.9	0.06	0.80	45.9 ± 2.77	0.84	0.36
cu	52.0 ± 5.8	0.29	0.87	50.0 ± 0.00	0.00	0.98
gl_1	19.1 ± 9.1	0.11	0.74	52.5 ± 2.07	0.60	0.44
fg	23.4 ± 9.0	1.31	0.25	48.8 ± 1.50	0.06	0.80
cl_1	18.5 ± 9.3	0.02	0.90	62.0 ± 4.74	8.87	0.00
R_1	48.6 ± 5.7	9.33	0.01	54.7 ± 2.83	1.16	0.28
$L_2{}^o$	52.2 ± 5.7	4.84	0.09	50.4 ± 0.80	0.06	0.80
T_1	44.2 ± 7.4	0.49	0.48	53.8 ± 2.41	0.36	0.55
R_2	43.0 ± 7.5	0.13	0.72	46.6 ± 2.41	0.84	0.36
P_1	54.9 ± 6.5	0.06	0.80	55.2 ± 2.98	1.81	0.18
Y_1	48.1 ± 7.1	2.81	0.09	44.8 ± 2.98	0.79	0.37
Lc_1	46.0 ± 7.3	3.27	0.07	53.4 ± 2.41	0.46	0.50
Lg	42.9 ± 7.5	1.75	0.19	48.3 ± 1.70	0.16	0.69
N_1	36.7 ± 7.9	7.85	0.01	20.7 ± 7.08	41.28	0.00

† Ot_4 and Ot_6 = populations derived from respective allele.
‡ RC% ± SE = Recombination percent ± standard error.
§ G = approximation to Chi Square, P > G = probability of a greater G value.

The results of the linkage tests contained the usual random linkage deviations, but there were consistent linkages of both transgenic loci with N_1, which was the only marker on chromosome 12 (Table 2). We hoped that we might find a linkage, but we thought it unlikely to find linkage associations with both transgenic loci. However, we observed segregation that indicated that both insertions were linked to N_1 (18 - 37 cM). The distances were slightly different, suggesting that they were not at the same site but that the insertions occurred in a narrow region of the chromosome. The data were consistent in that one transgenic locus showed apparent tighter linkage with N_1 than the other; however, the differences were not large and not statistically significant when data were pooled across the two transgenic populations.

These data establish linkage of Ot_4 and Ot_6 with N_1. The two-point linkage data cannot distinguish the orientation of the two inserts with respect to the N_1 locus. The two transformants were crossed to test the co-segregation or recombination between sites. The F_2 ($Ot_4 \times Ot_6$) was treated with 2,4-D, and all 2132 F_2 seedlings were resistant to 2,4-D and the 34 controls were susceptible (Fig. 1).

Table 3. Segregation and analysis of progeny from telosomics for the long arm of chromosome 12 crossed with transgenic cottons resistant to 2,4-D (Te12Lo's × Ot_n)F_2's.

Cross†	Segregation (*n*) Resistant	Susceptible	X^2 (3:1)	P
($Te^h \times Ot_4$)F_2				
disomic F_1	197	62	0.11	0.80–0.70
aneuploid F_1	36	6	2.57	0.10–0.05
($Te^b \times Ot_4$)F_2				
disomic F_1	323	85	3.52	0.10–0.05
aneuploid F_1	36	2	7.89	< 0.01
($Te^h \times Ot_6$)F_2				
disomic F_1	274	63	6.39	0.02–0.01
aneuploid F_1	381	57	33.56	< 0.01
($Te^b \times Ot_6$)F_2				
disomic F_1	239	74	0.31	0.70–0.50
aneuploid F_1	73	6	12.76	< 0.01

† = Te^h = (Te12Lo), Te^b = (Te12Lo × 3-79).

Crosses were made with monotelodisomic lines missing the short arm of chromosome 12 to establish arm location on chromosome 12. The monotelodisomic lines Te12Lo, isogenic with TM 1, and Te12Lo × 3-79, the *G. barbadense* F_1 cross, were crossed with Ot_4 and Ot_6. The F_1 were scored for aneuploid vs. wild-type phenotypes, and each F_2 progeny was tested for resistance vs. susceptibility to 2,4-D. Progeny of both the wild-type and aneuploid phenotype F_1's progeny segregated for resistance vs. susceptible (Table 3). This result established that both Ot_4 and Ot_6 were located in the long arm of chromosome 12.

With our original observation of linkage in both lines with N_1, we reviewed our lab notes on the original regeneration. The records indicate that these two lines originated from separate hypocotyl segments, but it is extremely difficult to prove unequivocally that an experimental error did not occur. We have conducted molecular analyses of these lines, but the results have not satisfied critical review to prove that the lines are different. The results presented have identified an *Ot* gene as a new marker on chromosome 12. It has been established that independent multiple insertions do occur (Bayley et al., 1992). Despite our interest in whether there are preferred sites for gene insertion, and because of the obvious impact on breeding of transgenic cottons, these questions will have to remain for other research

as we do not have access to transgenic materials to adequately address this question. This report will have to limit its findings to establishing at least one new member of the linkage group V on chromosome 12.

REFERENCES

Assaad, F.F., and E.R. Signer. 1992. Somatic and germinal recombination of a direct repeat in *Arabidopsis*. Genetics 132:553–566.

Bayley, C., N. Trolinder, C. Ray, M. Morgan, J.E. Quisenberry, and D.W. Ow. 1992. Engineering 2,4-D resistance into cotton. Theor. Appl. Genet. 83:645–649.

Endrizzi, J.E., E.L. Turcotte, and R.J. Kohel. 1984. Qualitative genetics, cytology, and cytogenetics. p. 81–129. *In*. R. J. Kohel and C. F. Lewis (ed.) Cotton. ASA, CSSA, and SSA, Madison, WI.

Firoozabady, E., D.L. DeBoer, D.J. Merlo, E.L. Halk, L.N. Amerson, K.E. Rashka, and E.E. Murray. 1987. Transformation of cotton (*Gossypium hirsutum* L.) by *Agrobacterium tumefaciens* and regeneration of transgenic plants. Plant. Mol. Biol. 10:105–116.

Liu, B.H., and S.J. Knapp. 1991. A reference manual for G-MENDEL 1.0: Statistical software for Mendelian segregation and linkage analyses. Oregon State Univ. Corvallis, OR.

Matzke, M.A., F. Neuhuber, and A.J.M. Matzke. 1993. A variety of epistatic interactions can occur between partially homologous transgene loci brought together by sexual crossing. Mol. Gen. Genet. 236:379–386.

Misra, S. 1990. Transformation of *Brassica napus* L. with a 'disarmed' octopine plasmid of *Agrobacterium tumefaciens*: Molecular analysis and inheritance of the transformed phenotype. J. Exp. Bot. 41:269–275.

Misra, S., and L. Gedamu. 1989. Heavy metal tolerant transgenic *Brassica napus* L. and *Nicotiana tabacum* L. plants. Theor. Appl. Genet. 78:161–168.

Puonti-Kaerlas, J., T. Eriksson, and P. Engstrom. 1992. Inheritance of a bacterial hygromycin phosphotransferase gene in the progeny of primary transgenic pea plants. Theor. Appl. Genet. 84:443–450.

Scheid, O.M., J. Paszkowski, and I. Potrykus. 1991. Reversible inactivation of a transgene in *Arabidopsis thaliana*. Mol. Gen. Genet. 228:104–112.

Umbeck, P., G. Johnson, K. Barton, and W. Swain. 1987. Genetically transformed cotton (*Gossypium hirsutum* L.) plants. Bio/Technology 5:263–266.