QTL Analysis of Transgressive Segregation in an Interspecific Tomato Cross

M. C. deVicente and S. D. Tanksley

Department of Plant Breeding and Biometry, Cornell University, Ithaca, New York 14853

Manuscript received April 29, 1992
Accepted for publication March 2, 1993

ABSTRACT

Two accessions, representing the species *Lycopersicon esculentum* (cultivated tomato) and *Lycopersicon pennellii* (a wild relative), were evaluated for 11 quantitative traits and found to be significantly different for 10 of the traits. Transgressive segregation was observed for eight of the traits in a large interspecific F_2 population. When restriction fragment length polymorphism markers were used as probes for the quantitative trait loci (QTL) underlying the traits, 74 significant QTL (LOD > 2) were detected. Thirty-six percent of those QTL had alleles with effects opposite to those predicted by the parental phenotypes. These QTL were directly related to the appearance of transgressive individuals in the F_2 for those traits which showed transgressive segregration. However, the same types of QTL (with allelic effects opposite to those predicted by the parents) were also observed for traits that did not display transgressive segregation in the F_2. One such trait was dry weight accumulation. When two overdominant QTL (detected in the F_2) for this trait were backcrossed into the *L. esculentum* genetic background, transgressive individuals were recovered and their occurrence was associated with the two QTL demonstrating the potential for transgressive segregation for all characters and implicating overdominance as a second cause of transgressive segregation. Epistasis was not implicated in transgressive segregation in either the F_2 or backcross generations. Results from this research not only reveal the basis of wide-cross transgressive segregation, but demonstrate that molecular markers can be used to identify QTL (from wild species) responsible for transgressive phenotypes and to selectively transfer them into crop species. This strategy might be used to improve many traits of economic importance including those for which wild species appear phenotypically inferior to their cultivated counterparts.

TRANSGRESSION is defined as the appearance of individuals in segregating populations that fall beyond their parental phenotypes. It is often observed in the progeny derived from interspecific matings (DARLINGTON and MATHER 1949; LOTSY 1916; BRAINERD 1924; VEGA and FREY 1980). RICK and SMITH (1953) proposed three explanations for the occurrence of interspecific transgression: (1) *de novo* mutation induced by the wide cross itself, (2) complementary action of genes from the two parental species, and (3) unmasking of recessive genes normally held heterozygous in the wild species. *De novo* mutation due to activation of quiescent transposable elements has been demonstrated between strains of Drosophila, but has not been demonstrated in plants (ENGELS 1983). However, evidence supporting both of the latter two hypotheses can be found in experiments involving a variety of plant species (RICK and SMITH 1953; FREY *et al.* 1981; VEGA and FREY 1980).

Interspecific transgression has evolutionary implications since it can potentially affect characters of adaptive significance leading to new races or species (STEBBINS 1950; LEWONTIN and BIRCH 1966). Moreover, transgressive individuals may possess characteristics that will allow them to occupy new ecological niches or to better compete in existing environments (LEWONTIN and BIRCH 1966). Interspecific transgression is also significant with respect to crop improvement since it represents a potential source of novel genetic variation. While many of the transgressive phenotypes observed in progeny of wide crosses are of no obvious practical value, there are a number of documented cases where novel characters of agronomic importance have appeared in offspring from the mating of crop plants with their wild relatives (FREY *et al.* 1981; RICK and SMITH 1953; REEVES and BOCKHOLT 1964; HARLAN 1976). Gaining a better understanding of the genetic basis of wide-cross transgression is therefore not only desirable from an academic view point, but also because of its practical implications. If one could learn to predict and manipulate this type of genetic variation, it might be harnessed to produce crop varieties with novel and/or exceptional characteristics.

The current study was an attempt to answer three questions regarding interspecific transgression. (1) Does transgression occur for most or only some of the characters segregating in wide crosses? (2) What fac-

tors predict the occurrence and extent of transgressive variation? (3) Can the individual quantitative trait loci (QTL), responsible for transgressive phenotypes, be identified and mapped (with molecular markers) and if so what is the nature of their gene action? The tools for conducting this study were a wide cross in tomato *Lycopersicon esculentum* × *Lycopersicon pennellii* (the former is the cultivated tomato and the latter is a distantly related, but crossable, species from Peru) and a well populated restriction fragment length polymorphism (RFLP) map for the tomato genome (TANKSLEY *et al.* 1992).

MATERIALS AND METHODS

Plant material: *L. esculentum* cv Vendor TM2a (hereafter referred to as *E*), was crossed to *L. pennellii* LA716 (hereafter referred to as *P*) using *E* as the female. The F_1 was self-pollinated to produce F_2 seed which were subsequently germinated in Petri dishes and transferred to flats in the greenhouse. Nine days after the appearance of the first true leaf, 432 F_2 seedlings were tranferred to individual 15-cm diameter pots. Control plants (32 *E* plants; 18 *P* plants, and 33 F_1 plants), treated the same way, were also included in the experiment. Pots were completely randomized and rotated weekly to minimize environmental effects due to heterogeneity in the greenhouse. The average day length was 14.5 hr and the average daytime temperature was 30°.

The same F_1 interspecific hybrid used to generate the F_2 population was backcrossed to *E* using the latter as the female. A single BC1 individual, selected on the basis of its RFLP genotype in chromosomes *1* and *2* (see Figure 1), was crossed again to *E* to produce a BC2 population. BC2 seeds were sown in the greenhouse in May 1991, and 142 BC2 plants were transplanted to the field in a completely randomized design in Ithaca on June 4 along with 15 F_1 and *E* controls.

A single BC2 plant was selected from among the transgressive individuals in the field experiment of 1991 and the selfed progeny of that plant were grown in the greenhouse and analyzed with RFLPs. A single BC2F2 plant was then selected on the basis of its genotype. One hundred forty-eight BC2F3 progeny from this plant, along with 17 *E* controls, were planted again in a different field in Ithaca on May 29, 1992, in a completely randomized experiment.

Characters: After transfer to the greenhouse, individual F_2 plants were monitored for appearance of the first true leaf (days to first true leaf = DTL) and the first open flower (days to first flower = DTF). Forty-five days after the appearance of the first true leaf, the following additional characters were recorded for each individual: plant height (HT), total number of flower buds (BUDS), number of internodes on the primary stem (stem nodes = SNOD), total number of internodes (TNOD), number of well-developed branches (BRAN), total fresh weight (FW) and total dry weight (DW) of the aerial portion of the plant, diameter of the stem at the first internode (DIAM) and leaflet width/length ratio (LR). LR was determined by averaging measurements from three separate leaflets from each plant. To obtain FW and DW measures, plants were cut 10 cm above the ground and allowed to regrow for DNA isolation.

Means, standard errors and standard deviations were determined for each character for the two parents, F_1 hybrid and F_2 population using Super ANOVA v1.1 (Abacus Concepts). Normality of the F_2 distribution for each character was tested with normal probability plots. To improve normality, data were transformed for DTL ($\log_{10}$) and BUDS (square root). Tests for significant differences in means for each character among *E*, *P*, and the F_1 were made using contrasts from analysis of variance.

RFLP analysis: Ninety-eight RFLP markers, chosen to cover the tomato genome at intervals of approximately 10–20 cM, were used to construct the linkage map of the F_2 population, which served as the basis for QTL analysis. The BC2 field population was analyzed for 42 markers from selected regions of the genome (Figure 1). The F_2-derived map was constructed with MAPMAKER computer software using the Kosambi map unit function (LANDER *et al.* 1987; KOSAMBI 1944). Procedures for DNA extraction, restriction enzyme digests and Southern blotting were as described in BERNATZKY and TANKSLEY (1986). Labeling of probes was via primer extension (FEINBURG and VOGELSTEIN 1983). HYPERGENE (YOUNG and TANKSLEY 1989) was used to construct graphical genotypes, calculate genome ratios (percentage of total genome comprised of one parental genome) and calculate percent homozygosity and heterozygosity for each plant in the segregating population. The homozygosity and heterozygosity percentages were estimated based on the genotypes of the consecutive markers along a particular chromosome. If the individual has the same genotype at two of those markers, the chromosome region between them is also considered to have that genotype and to be homozygous. On the contrary, if two consecutive markers have different genotype, then that interval would be half of each genotype and considered heterozygous.

QTL mapping: The association of markers with the QTL for each trait was tested using two different procedures. The first was a single point analysis using a one-way ANOVA from the PROC GLM routine in SAS (SAS Institute, 1988) with each marker considered as a treatment. This analysis included contrasts to compare, for each RFLP marker and trait, the genotypic means of the homozygous and heterozygous classes. The second was interval mapping which searches for the effects of QTL using sets of linked markers (LANDER and BOTSTEIN 1989). A LOD score of 2.0 was used as the threshold for detecting QTL locations. LOD peaks for each significant QTL were used to position QTL on the RFLP linkage map. The additive effect (a), the dominance deviation (d) and the degree of dominance (d/a) were calculated for each QTL by MAPMAKER QTL software. MAPMAKER QTL was also used to test the gene action of each QTL against the dominant, additive, recessive, overdominant/underdominant (referred to as "unconstrained genetics" in MAPMAKER QTL models) (PATERSON *et al.* 1991). QTL with $d/a < $ or > 1 and which fit the unconstrained genetics model with LOD > 1 (compared with alternative models) were considered to be significantly overdominant/underdominant. Tests for two-way interactions were made between significant QTL and all other marker loci segregating in the population using the PROC GLM routine in SAS (SAS Institute, 1988).

Environmental variance for each trait was estimated by averaging the variance of the control plants (*E*, *P* and F_1). The percentage of phenotypic variation accounted for by all significant QTL for each trait in the F_2 was estimated from the multilocus model using MAPMAKER-QTL software. The percentage of genetic variance was estimated by multiplying the total variance of the F_2 (for each trait) by the estimated percentage of phenotypic variation and dividing this by the estimated genetic variance of the F_2 (total F_2 variance minus environmental error).

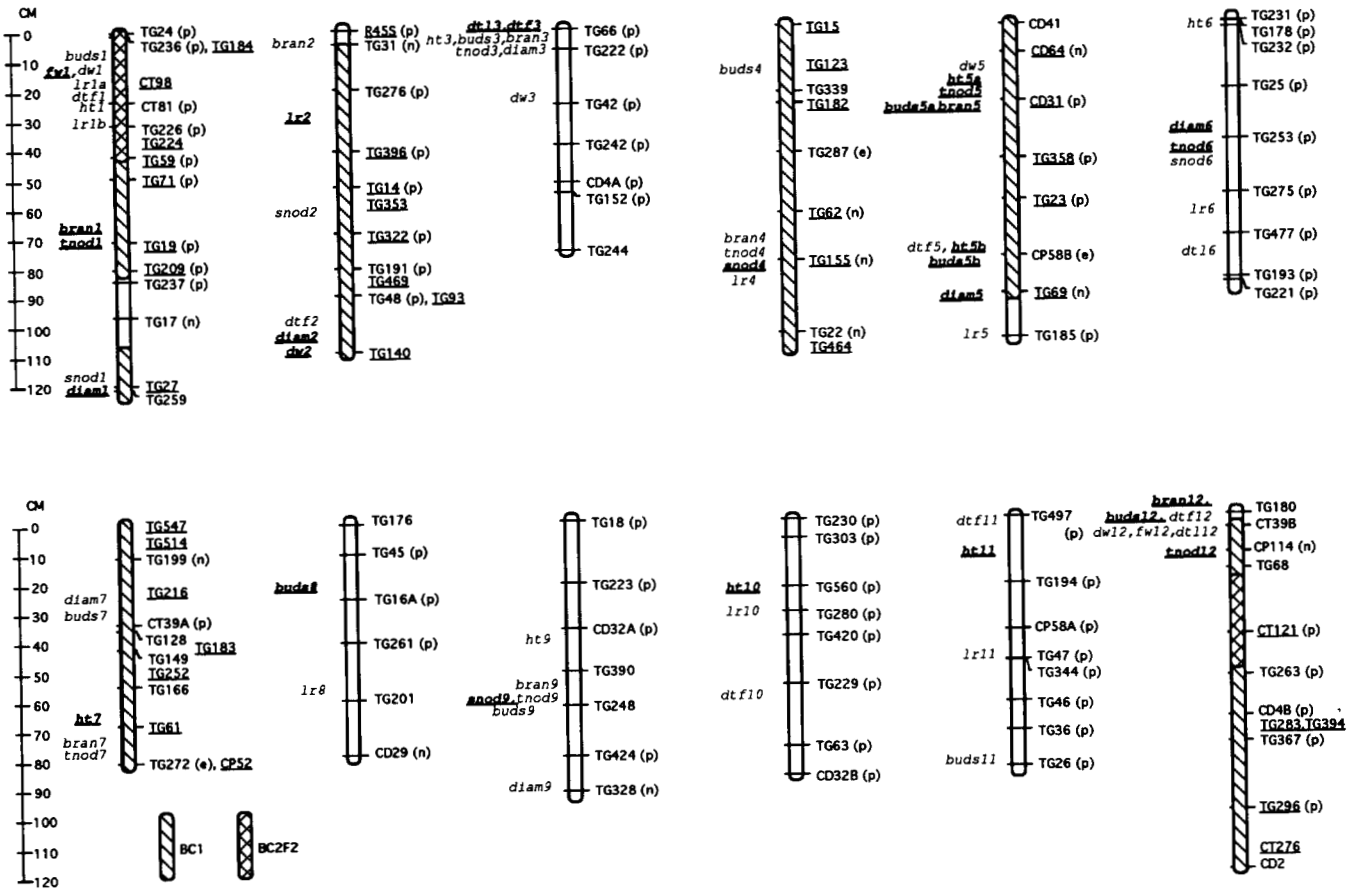

FIGURE 1.—Genetic linkage map of tomato derived from *L. esculentum* cv Vendor *Tm2a* × *L. pennellii* LA716 F_2 population of 432 F_2 plants. Symbols to right of chromosomes are RFLP-detected loci. (*p* = segregation signficantly skewed in favor of *P* allele. (*e*) = segregation signficantly skewed in favor of *E* allele. (*n*) = significantly skewed toward (or against) heterozygotes. Underlined loci were also scored in BC2 population (see text for details). Symbols to left of chromosomes represent approximate positions of significant QTL (LOD > 2) detected for 11 characters in the F_2 population (see Table 2 for details). QTL underlined and bold had allelic effects opposite of that predicted by the parental values. Shaded and unshaded chromosome segments correspond to those from *L. pennelli* and *L. esculentum*, respectively, in a single BC1 plant used to generate a BC2 population used for further evaluation of the effects of *dw1* and *dw2*.

RESULTS

RFLP segregation and map construction: The RFLP linkage map of the F_2 covered 1141 cM and the average distance between markers was 13.5 cM (Figure 1). The order of all markers on the map was the same as that previously reported (DEVICENTE and TANKSLEY 1991; TANKSLEY *et al.* 1992). Seventy-eight (80%) of the 98 marker loci analyzed deviated significantly ($P < 0.05$) from the expected 1:2:1 frequency. These loci represent 16 linked segments spread around all 12 chromosomes (Figure 1). Of those 78 markers, 64 (82%) showed significant deviation toward the *P* allele, while only 3 loci (4%) were skewed toward the *E* allele. Results from the F_2 are consistent with other studies using the same interspecific cross in tomato and in which deviations toward the *P* allele in the F_2 have been also reported (ZAMIR and TADMOR 1986). There was no evident trend of the distorted segregations being related to the appearance of transgressive QTLs (Figure 1).

Genome composition: The average percent heterozygosity per plant in the F_2 population was 52 ± 8.7%, with a range from 22% to 87%. The average heterozygosity for an F_2 individual was close to the expected (50%), and the range exceeded three standard deviations on each side. The average percent homozygosity was 17 ± 6.9% (range: 0–41%) for *E*/*E* genotypes and 32 ± 9.5% (range: 7–94%) for *P*/*P* genotypes. The parental genome ratio was 43% *E*: 57% *P*. Deviations from expected homozygous ratios were observed with an overabundance of *P* alleles in most cases.

Character evaluations: The cultivated tomato (*E*) and its wild relative *L. pennellii* (*P*) differed significantly for 10 of the 11 traits examined (Table 1). Only for HT (height) were the parental means indistinguishable (Table 1). In all cases, F_2 progeny were observed that fell beyond the high or low mean of the two parents (Figure 2). The occurrence of some extreme F_2 individuals is expected due to environmental rather than genetic reasons since the F_2 population was much larger than the parental populations. Chi-

TABLE 1

Controls means and standard deviations for all traits studied in the F_2 experiment

Trait	*E*	*P*	F_1	F_2	Contrasts *E/P*	*MP/F*$_1$	*E/F*$_1$	F_1/P
DTL	10.3 ± 0.1	(11.7 ± 0.2)	18 ± 0.4	11.7 ± 2.3	**	**	**	**
DTF	33.5 ± 0.3	(45.0 ± 0.0)	37 ± 0.3	39.1 ± 3.4	**	*	**	**
HT	54.9 ± 0.7	57.1 ± 1.5	70.4 ± 2.3	66.1 ± 16.3	NS	**	**	**
BUDS	22.9 ± 1.0	(55.6 ± 3.6)	40 ± 2.4	34.6 ± 20.5	**	NS	**	**
SNOD	(13.8 ± 0.2)	12.9 ± 0.3	14.1 ± 0.2	13.8 ± 1.2	*	*	NS	*
TNOD	21.6 ± 0.5	(28.2 ± 1.5)	30.4 ± 1.3	27.8 ± 10.0	**	**	**	NS
BRAN	2.7 ± 0.1	(3.4 ± 0.3)	4.6 ± 0.2	3.7 ± 1.7	*	**	**	**
FW	(108.5 ± 3.8)	56.0 ± 3.4	97.4 ± 4.2	83.8 ± 25.7	**	*	*	**
DW	(19.3 ± 0.6)	10.0 ± 0.6	20 ± 0.7	15.1 ± 3.6	**	**	NS	**
DIAM	(10.8 ± 0.2)	8.1 ± 0.3	10.6 ± 0.3	10.6 ± 2.0	**	**	NS	**
LR	0.6 ± 0.0	(1.0 ± 0.0)	0.7 ± 0.0	0.7 ± 0.1	**	**	**	**

Significance for the contrasts (comparisons among means) is represented with asterisks: $P < 0.05$ (*), $P < 0.001$ (**), NS = non-significant. *E* = *esculentum*, *P* = *pennellii*, F_1 = interspecific hybrid ($E \times P$), F_2 = self-progeny of the interspecific hybrid, *MP* = mid-parent value. Parentheses are used to designate the control parent with a significant higher value for each particular trait. Underlining is used to show cases in which the F_1 was significantly greater than both parents.

square tests were used to compare the number of individuals (in a population the size of the F_2) expected to exceed the means of the parents by at least 2 SD (observed in those parental distributions and assuming that the size of those populations were equal to the F_2) with the actual number of F_2 individuals that exceeded those thresholds. For all characters, except DW, FW and LR, the number of observed extreme individuals significantly ($P < 0.01$) exceeded the expected, suggesting transgressive segregation. In some instances transgression was extreme (*e.g.*, HT, TNOD, DTL) with some F_2 individuals having values more than two fold higher or lower than the parental values (Figure 2). For DW, FW and LR, the number of extreme individuals was actually less than expected due to environmental effects alone (data not shown).

Relationship between heterozygosity and expression of quantitative characters: The effect of heterozygosity on each trait was estimated by regressing the phenotypic value of each individual F_2 plant on its percent heterozygosity (determined by RFLP probing, see MATERIALS AND METHODS for details). Values obtained from such regression analyses can be taken as an estimation of the effect of heterozygosity *per se* on a particular trait (EDWARDS, STUBER and WENDEL 1987). *R*-squared values ranged from 0.0% to 4.3% and only for DTL, DTF, DW and LR were these values significantly different from zero ($P < 0.05$). These results indicate that only a small portion of the overall variability in the F_2 can be explained by heterozygosity alone.

QTL affecting characters: Each character was subjected to QTL mapping using RFLP markers mapped in the F_2 population. A total of 74 QTL (LOD > 2) were detected for all characters with a range of 2–10 QTLs detected per character (Figure 1, Table 2). The percentage of phenotypic variance explained by each QTL for each character ranged from 2.6% (for DIAM) to 34.0% (for BUDS) (Table 2). Multilocus analysis indicated that the combined action of *all* the QTL detected for each character could account for 7–60% of the phenotypic variation and 15–88% of the genotypic variation in the F_2 (Table 3). The phenotypic and genotypic variation attributable to detected QTL, averaged over all characters, was 37% and 56%, respectively (Table 3).

Epistasis: Markers linked to each of the 74 significant QTL were tested for possible two-way interactions with all other segregating markers (total of 7081 two-way tests). Three hundred and forty-six (4.9%) of the interactions were significant ($P \leq 0.5$) which is close to the percentage expected by chance. The loci showing significant interactions seldom coincided with the location of other QTL for the same character suggesting that epistasis was not a major factor underlying the genetic inheritance of these traits.

Gene action of QTL: A histogram of the d/a values for all QTL indicates a continuum of gene action centered approximately at zero and ranging from underdominance ($d/a < -1.0$) to overdominance ($d/a > 1.0$) (Figure 3). Overall, 7% (5/74) of the QTL showed significant overdominance/underdominance. A few QTLs (*e.g.*, for DTF, BUDS and DW) showed extreme overdominance/underdominance and will be discussed later (Table 2).

By comparing the difference between the means of the parental controls with the genotypic (*a*) value of each QTL allele, it was possible to determine whether each allele had an effect in the expected or opposite direction relative to the parental means. For example, if *P* had a value greater than *E*, any QTL allele from *P* that increased the character in the F_2 would be in the expected direction and any QTL allele from *P* that decreased the character would be in the opposite

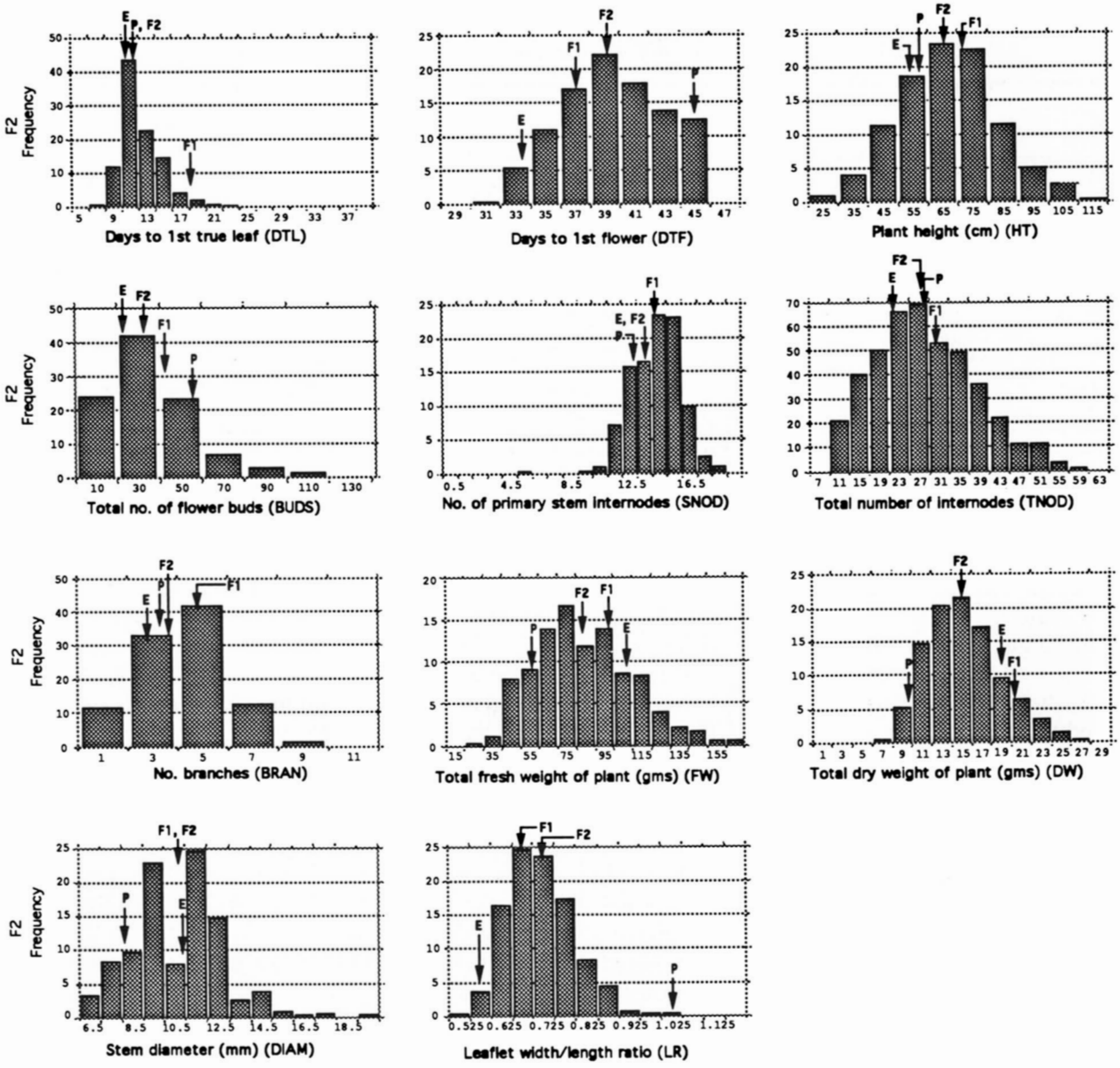

FIGURE 2.—Distribution of phenotypes for each character in the F_2 population. Means for parental (*E* and *P*) and F_1 controls are shown by arrows.

direction. Using this method of classification, at least one QTL for each character was found to have allelic effects in the opposite direction from that predicted by the parental phenotypes, and overall 36% of the QTL were of this nature (Figure 4).

Relationship between gene action of QTL and transgressive phenotypes: QTL with effects opposite to those predicted by the parents or with overdominance/underdominance could be directly responsible for the occurrence of individuals with transgressive phenotypes in the F_2 population. For example, transgressive individuals might have a combination of complementary positive QTL alleles from both parents, causing them to exceed the parental phenotypes. This hypothesis predicts that characters for which many transgressive individuals appear are also likely to be the characters for which a large proportion of the QTL have allelic effects opposite of that predicted by the parents. To test this hypothesis, a logistic regression was run between the percentage of QTL with opposite allelic effects and the number of individuals which surpassed the parental means by at least two standard deviations (both high and low). The results suggest a strong positive relationship between these two variables (Table 4). To confirm this, three F_2 plants were selected from the high end of the distribution for three characters: HT, TNOD and DIAM. These traits were ones in which extreme transgression had been observed. Graphical genotypes (YOUNG and TANKSLEY 1989) were generated for each plant and used to examine the allelic composition at the significant QTL for each character. For HT, on average, the tallest plants had a combination of 2.3 QTL alleles increasing height (1 QTL from *E* and 1.3 QTL from

TABLE 2

List of QTLs detected for each trait

Trait	QTL	LOD	Percent variance	*a*	*d*	*d*/[*a*]
DTL	(*dtl3*)	11.91	13.1	0.04	−0.04 (p)	−1.10
	dtl6	2.72	3.8	−0.02	−0.02 (e)	−1.23
	dtl12	2.43	2.6	−0.02	−0.01 (e)	−0.88
DTF	*dtf1*	3.11	8.0	−0.04	−1.91 (p)	−51.87
	dtf2	2.48	5.5	−0.08	−1.58 (e)	−18.96
	(*dtf3*)	8.00	10.2	1.96	−0.47 (p)	−0.24
	dtf5	6.93	12.2	−1.67	0.80 (p)	0.48
	dtf10	2.20	4.0	−0.84	−0.48 (e)	−0.58
	dtf11	2.50	3.5	−0.71	1.17 (p)	1.65
	dtf12	3.34	4.4	−0.83	−0.78 (e)	−0.94
HT	*ht1*	4.49	7.2	−5.01	7.55 (p)	1.51
	ht3	4.35	5.2	−4.83	6.16 (p)	1.28
	(*ht5a*)	3.85	5.0	4.60	2.47 (e)	0.54
	(*ht5b*)	6.05	8.2	7.53	−0.86 (p)	−0.11
	ht6	8.20	8.4	−7.63	1.44 (p)	0.19
	(*ht7*)	2.39	3.3	4.41	2.42 (e)	0.55
	ht9	2.38	3.1	−3.61	3.05 (p)	0.85
	(*ht10*)	4.54	5.6	7.97	−6.77 (p)	−0.85
	(*ht11*)	2.22	4.0	3.85	2.87 (e)	0.74
BUDS	*buds1*	2.76	6.6	−0.01	0.91 (p)	79.39
	buds3	34.11	34.0	−1.61	0.65 (p)	0.40
	buds4	2.78	4.2	−0.45	−0.35 (e)	−0.79
	(*buds5a*)	4.17	4.6	0.48	−0.47 (p)	−0.98
	(*buds5b*)	5.30	5.6	0.57	−0.44 (p)	−0.78
	buds7	3.03	3.3	−0.48	0.07 (n)	0.14
	(*buds8*)	2.20	3.2	0.47	−0.23 (p)	−0.48
	buds9	4.95	5.8	−0.59	−0.18 (e)	−0.31
	buds11	3.19	3.5	−0.32	−0.42 (e)	−1.30
	(*buds12*)	2.54	2.8	0.38	−0.27 (p)	−0.71
SNOD	*snod1*	4.35	6.5	0.62	−0.09 (p)	−0.14
	snod2	5.39	7.4	0.64	−0.28 (p)	−0.43
	(*snod4*)	3.35	4.9	−0.49	0.42 (p)	0.86
	snod6	3.18	4.8	0.71	−0.14 (p)	−0.20
	(*snod9*)	5.71	6.4	−0.60	0.03 (n)	0.05
TNOD	(*tnod1*)	5.08	6.5	3.16	1.48 (e)	0.47
	tnod2	2.75	4.7	−2.40	−1.53 (p)	−0.64
	tnod3	7.98	8.9	−4.70	1.80 (p)	0.38
	tnod4	4.77	6.2	−3.62	1.31 (p)	0.36
	(*tnod5*)	3.66	4.7	2.59	−2.77 (p)	−1.07
	(*tnod6*)	3.01	3.7	3.34	−0.46 (p)	−0.14
	tnod7	3.51	4.6	−2.89	1.73 (p)	0.60
	tnod9	10.76	12.8	−5.12	−0.26 (e)	−0.05
	(*tnod12*)	4.25	5.0	2.79	−2.23 (p)	−0.80
BRAN	(*bran1*)	5.27	8.2	0.71	0.02 (n)	0.02
	bran2	3.67	9.5	−0.43	−0.74 (e)	−1.74
	bran3	5.00	6.0	−0.66	0.31 (p)	0.47
	bran4	4.68	6.6	−0.63	0.30 (p)	0.48
	(*bran5*)	3.44	4.1	0.46	−0.35 (p)	−0.75
	bran7	4.66	6.9	−0.38	0.68 (p)	1.79
	bran9	7.50	9.5	−0.74	0.16 (p)	0.21
	(*bran12*)	2.74	3.1	0.36	−0.32 (p)	−0.89
FW	(*fw1*)	2.19	4.4	−4.01	10.81 (p)	2.70
	fw12	2.74	3.0	6.01	2.15 (e)	0.36
DW	*dw1*	3.40	7.0	0.01	1.92 (e)	233.54
	(*dw2*)	2.33	3.7	−0.22	1.38 (p)	6.38
	dw3	2.51	3.2	0.78	0.52 (p)	0.67
	dw5	3.05	4.0	1.04	−0.09 (p)	−0.09
	dw12	3.48	3.8	0.95	0.38 (e)	0.40
DIAM	(*diam1*)	3.26	3.6	−0.37	−0.51 (e)	−1.40
	(*diam2*)	3.09	4.5	−0.57	0.40 (p)	0.70
	diam3	4.48	5.4	0.71	−0.06 (e)	−0.09
	(*diam5*)	4.14	4.4	−0.60	0.36 (p)	0.61
	(*diam6*)	4.37	6.3	−0.74	−0.17 (p)	−0.23
	diam7	3.07	4.9	0.63	0.25 (e)	0.39
	diam9	2.26	2.6	0.46	0.05 (e)	0.10
LR	*lr1a*	21.56	29.8	−0.06	−0.02 (e)	−0.41
	lr1b	19.64	21.2	−0.05	−0.02 (e)	−0.41
	(*lr2*)	2.68	4.3	0.02	−0.02 (p)	−0.73
	lr4	5.03	7.1	−0.03	0.00 (p)	0.04
	lr5	2.55	3.7	−0.01	−0.02 (e)	−1.50
	lr6	2.30	3.1	−0.02	−0.01 (e)	−0.33
	lr8	9.39	11.6	−0.04	0.00 (n)	0.08
	lr10	2.62	3.1	−0.02	−0.01 (e)	−0.32
	lr11	8.01	9.1	−0.02	−0.03 (e)	−1.18

The number following the QTL designation corresponds to the chromosome to which the QTL was located. Letters are used to distinguish QTLs on the same chromosome affecting the same trait. Parentheses (in the QTL column) indicate that the QTL had an effect opposite to that expected from the parental means. For each QTL: LOD = log-likelihood, percent variance = percent phenotypic variance explained (single-locus model), *a* = genotypic value = (*EE*-*PP*)/2, as the effect of adding an *E* allele, *d* = dominance effect = *EP*-(*EE* + *PP*)/2, and d/[*a*] = degree of dominance. The letter in parenthesis by the d value indicates the origin of the dominant allele (e = *esculentum*, p = *pennellii*, n = no dominance). Underlined *d*/[*a*] values show significant overdominance or underdominance (Δ LOD > 1, see MATERIALS AND METHODS for details).

P) and only 0.3 QTL decreasing height (from *P*). Similar results were seen for TNOD and DIAM, where the extreme individuals contained complementary combinations of QTL alleles from the two parents (data not shown).

Relationship between phenotypic diversity of parents and degree of transgression: A regression between the difference in the parental means for each trait and the number of transgressive individuals appearing in the F_2 revealed a significant negative association between these variables (Table 4). In other words the more similar the phenotype of the parent, the more likely that transgressive individuals would occur in the F_2. These results at first appear to be counterintuitive, but would be explained if the two species contain different mixtures of positive and negative QTL alleles leading to similar phenotypic means for some characters. If this is the case, one would predict a high proportion of QTL with allelic effects opposite to that predicted by the parents for characters for which the parents had similar means. This prediction was also borne out in a significant negative correlation between parental mean differences and the percentage of QTL with allelic effects opposite of

TABLE 3

Percentages of phenotypic and genetic variance explained by the QTLs (multiple locus model)

Trait	Phenotypic	Genetic
DTL	18	27
DTF	43	51
HT	42	54
BUDS	61	88
SNOD	23	41
TNOD	52	72
BRAN	53	88
FW	7	15
DW	21	70
DIAM	26	41
LR	60	73
X	37	56

See MATERIALS AND METHODS for details.

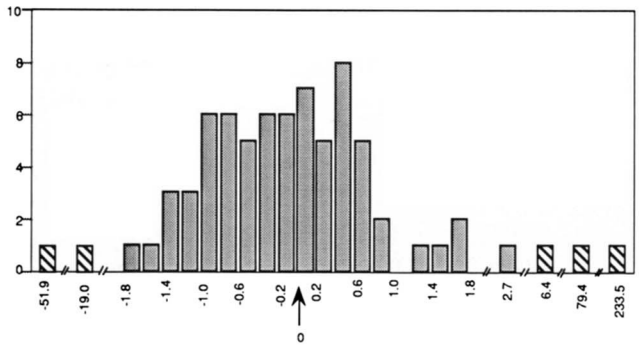

FIGURE 3.—Distribution of the degree of dominance (d/a) for individual QTL detected for the 11 traits studied in the F_2 generation. Striped bars correspond to QTL with significant overdominance ($d/a > 1.0$) or underdominance ($d/a < -1.0$).

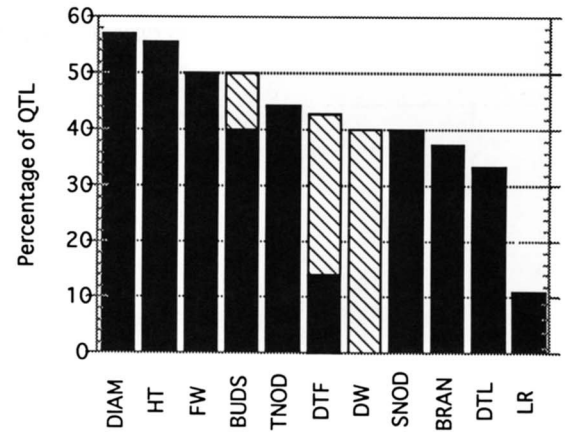

FIGURE 4.—Histogram depicting percentage of QTL detected with allelic effects opposite of that predicted by the parental phenotypes (solid) and QTL with overdominant gene action (striped). Both types of QTL can contribute to transgressive segregation (see text for details).

TABLE 4

Results from regression analyses of % transgressive QTL (% trans), number of transgressive progeny in F_2 (# trans) and distance between parental means (distance, expressed in standard deviations)

Explanatory variable	Response variable	Sign of the regression	Wald chi-square	$P >$ chi-square
% Trans	# Trans	+	102.6	0.0001
Distance	% Trans	−	4.5	0.0300
Distance	# trans	−	147.1	0.0001

All calculations based on F_2 data.

that predicted by the parental phenotypes (Table 4).

QTL with significant overdominance/underdominance behavior: Overdominance (or underdominance) potentially represents a single locus basis for trangressive variation and might account (at least in part) for some of the extreme phenotypes observed in the F_2 generation. However, the lack of a significant correlation in most instances between heterozygosity and phenotype variation in the F_2 (see previous section) suggests that overdominance was not likely to be a major factor in these experiments. Nonetheless, examination of the d/a values from Table 2 reveals that out of the 74 QTL detected in this study, 18 gave d/a values indicative of overdominance (or underdominance) ($d/a < -1$ or $d/a > +1$) 5 of which were significant (Table 2, Figure 3, see MATERIALS AND METHODS for details). Positive (overdominant) d/a values ($EP > EE$ and PP) were detected for BUDS and DW. Negative (underdominant) d/a values ($EP < EE$ and PP) were found for DTF. In some cases the d/a values were very extreme (*e.g.*, *buds1*, $d/a = 79$; *dw1*, $d/a = 233$).

Relationship between overdominant/underdominant QTL and heterosis: For DTL, HT and BRAN, heterosis could be observed in F_1 hybrids (*i.e.*, the F_1 mean was significantly greater or less both parental means) (Table 1). If overdominant/underdominant QTL were to be uncovered, one might expect to find them for traits showing heterosis in the F_1 hybrid. However, few such QTL were found for these characters. Instead, significant overdominant/underdominant QTL were detected for characters not showing heterosis in the F_1 (*e.g.*, DTF, BUDS, DW) (Table 2). For HT, where the greatest heterosis was observed, none of the QTL were overdominant (Table 2). All showed some degree of dominance in favor of the allele generating the extra height, regardless from which parent they came–a result that could also explain the heterotic response for this character in the F_1 hybrid.

Advanced backcross generation analysis of two overdominant QTL: *L. pennellii* grows at a rate approximately one-half that of the *L. esculentum* parent (Table 1) and while most of the QTL detected from *P* slowed down the growth rate, *dw1* (chromosome *1*) and *dw2* (chromosome *2*) increased the rate of dry matter accumulation in heterozygous F_2 individuals. *dw1* acts in an extremely overdominant manner ($d/a = 233$) and *dw2* in a less extreme manner ($d/a = 6$) (Table 2). Using linked RFLP markers, *dw1* and *dw2*

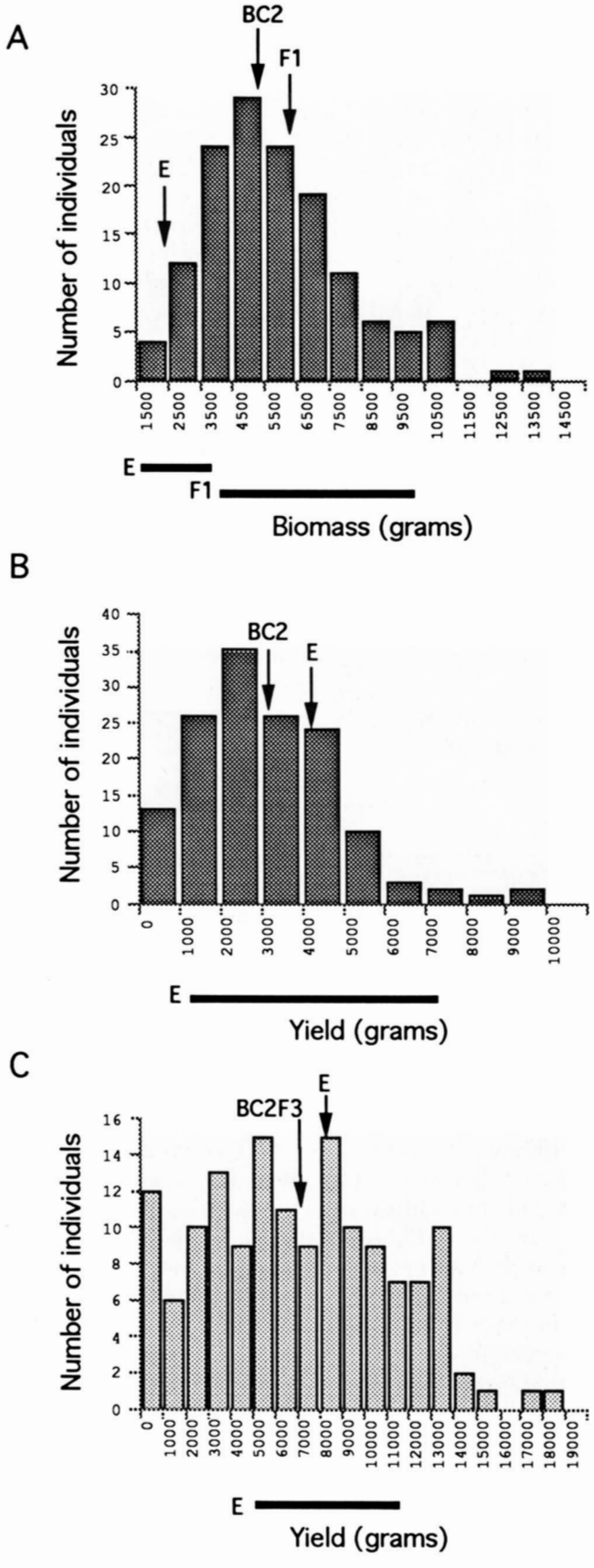

FIGURE 5.—Distribution of phenotypes in the BC2 population for dry weight accumulation (A) and fruit yield (B) and in the BC2F3 population for fruit yield (C). Arrows indicate means for *E* parent, F_1, BC2 and BC2F3 populations. Bars below plots represent the range observed for the *E* and F_1 controls. Fruit yield was not measured on F_1s because of low fruit set.

were transferred into an *L. esculentum* background via two sequential backcrosses to *E*. A BC2 population of 142 plants was transplanted to the field in Ithaca in the summer of 1991 along with 15 F_1 and 15 *E* control plants. *P* does not survive well under field conditions and was not utilized in these field experiments. The graphical genotype of the BC1 plant used to generate the BC2 population is shown in Figure 1. While additional *P* chromosomal segments (other than those containing *dw1* and *dw2*) were present, the majority of the chromosome segments (70.6%) in the BC1 individual were derived from *E* based on RFLP probing (Figure 1). The average BC2 individual in that population would be expected to contain only 15% *P* alleles and thus this population should provide a test of behavior of *dw1* and *dw2* in the absence of other potentially epistatic *P* alleles.

Approximately 7 weeks after transplanting to the field, each plant in the BC2 was rated visually for vigor. At the end of the season, the total biomass and fruit yield of each plant was determined. Total biomass and visual rating for vigor were significantly correlated ($R^2 = 0.3$, $P < 0.01$), and for subsequent analyses only biomass was used. A significant number of BC2 individuals exceeded the parental (*E*) and F_1 means (Figure 5A). RFLP markers linked to *dw1* and *dw2* were scored in the BC2 and used for QTL analysis of both biomass and fruit yield. RFLP markers linked to both *dw1* and *dw2* were significantly associated with increases in biomass accumulation and resulted in transgressive individuals (Figures 5 and 6). LOD scores and *P* values from ANOVA analysis were 1.4/0.02 and 4.7/0.0001 for *dw1*- and *dw2*-linked markers, respectively. Moreover, the gene action of each QTL was the same as that observed in the F_2 (*i.e.*, the heterozygous individual had increased biomass production). Although the LOD score was low, the ANOVA also indicated a significant, positive effect of *dw1* on fruit yield ($P = 0.05$) (Figure 6). *dw2* had no detectable effect on fruit yield (data not shown). These results indicate that both of the overdominant QTL detected in the F_2 continued to exert their effects in the BC2 where most of the other *P* alleles had been eliminated. Moreover, unlike with the F_2 generation, significant numbers of BC2 individuals were transgressive.

To further examine the relationship between *dw1* and yield, a single BC2F2 plant was selected from among the progeny of one of the transgressive BC2 plants from the field. This individual was heterozygous for the region on chromosome *1* containing *dw1*. Except for an additional small piece of *pennellii* DNA on chromosome *12*, this plant was determined to be

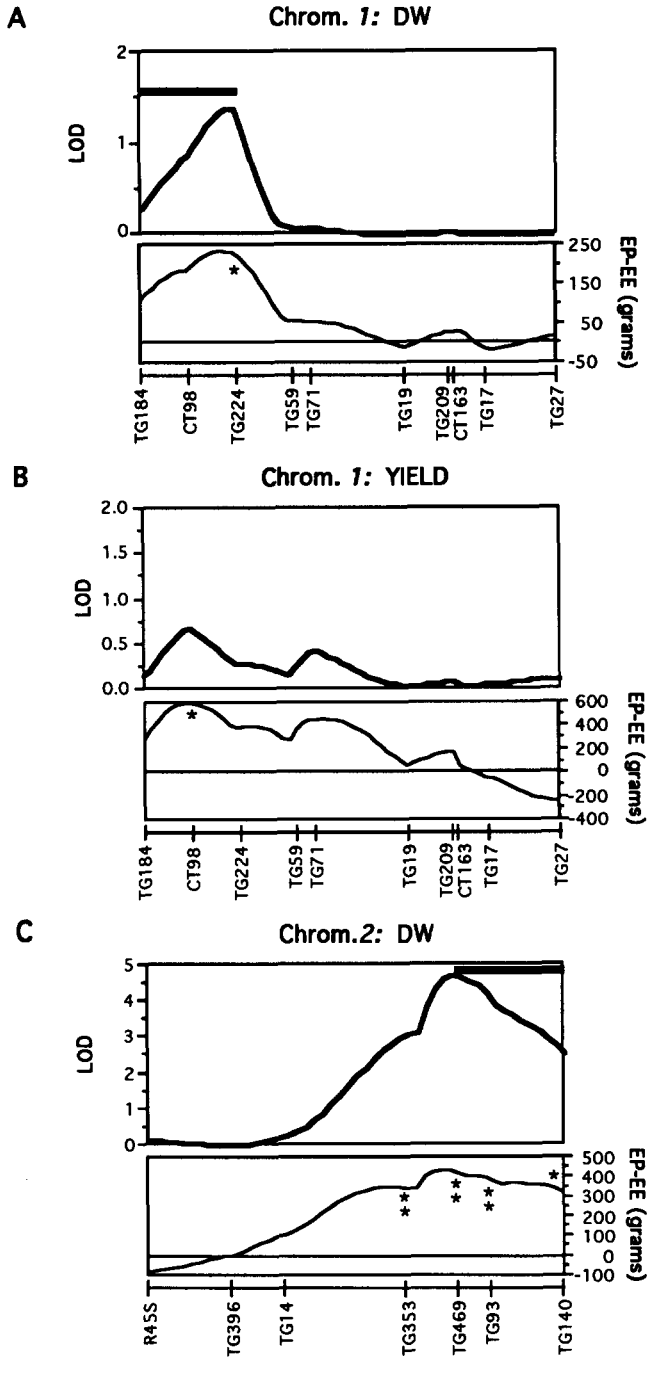

FIGURE 6.—QTL plots for *dw1* (chromosome *1*) and *dw2* (chromosome *2*) monitored in the BC2 population for biomass per plant (A, C) and fruit yield per plant (B). Upper half of each plot is for LOD scores based on MAPMAKER-QTL (see MATERIALS AND METHODS for details). Lower half is for difference between heterozygous plants (*E/P*) and homozygous plants (*E/E*) in the BC2 population, calculated with MAPMAKER-QTL at intervals of 2 cM. * = significant effect based on ANOVA ($P < 0.05$). Bar across top of plots indicates approximate position of *dw1* and *dw2* as deduced from F_2 analysis.

otherwise entirely of an *esculentum* genotype based on probing with RFLP markers throughout the genome (Figure 1). In the summer of 1992, a BC2F3 population of 148 plants, derived from self-pollinating this individual, as well as 17E controls, were transplanted to the field in Ithaca. Fifteen weeks later, total fruit yield/plant was determined for each individual. Twenty-six BC2F3 plant (18%) exceeded the controls (Figure 5C). The association of yield with RFLP markers in the *dw1* region was verified with a single point analysis of variance ($P = 0.001$). A one-tailed contrast revealed that individuals heterozygous at *dw1* significantly out yielded either homozygous genotype ($P < 0.03$), confirming that the gene action of this QTL is indeed overdominant as observed in previous populations. The genotypic means for the marker closest to *dw1* (CT98) were as follows: *EE* = 7.79 kg/plant, *EP* = 9.02 kg/plant, *PP* = 2.49 kg/plant.

Possible pleiotropic effects of *dw1* and *dw2*: In the F_2 population, two of the seven QTL affecting the number of days to flowering (*dtf1* and *dtf2*) mapped very close to *dw1* and *dw2*, respectively (Table 2, Figure 1). The *d/a* values for both of these QTL indicated a strong overdominance/underdominance gene action, a characteristic also displayed by *dw1* and *dw2*. The close proximity and similar gene action of these QTL controlling DW and DTF may be due to chance genetic linkage, but might also result from single QTL which affect both the growth rate and flowering date.

DISCUSSION

If different species carry a high frequency of complementary alleles, then interspecific hybridization has great potential for generating novel phenotypes–including those for characters important for adaptation in natural habitats as well as ones valuable in agriculture. Results from the current study suggest that this is likely to be the case. For eight of the characters examined, there was evidence for transgressive segregation based on the occurrence of significant numbers of extreme individuals in the F_2 population. The occurrence of this transgression was directly linked to the presence of complementary QTL alleles in the two parental species. In the cases where the parents were similar in their means (*e.g.*, plant height), progeny with extreme phenotypes were identified. For these same characters, a high proportion of QTL alleles with effects opposite of that predicted by the parents were identified and could be associated causally with progeny expressing transgressive phenotypes. This study therefore demonstrates that complementary QTL alleles occur in a wide cross in a frequency not dissimilar from intraspecific crosses (STUBER, EDWARDS and WENDEL 1987). Moreover, it establishes as the major cause of the transgression the complementary action of genes from the two parental species. It is important to note that for three of the characters examined, DW, FW and LR, there was no evidence of transgressive segregation in the F_2 popu-

lation. Nonetheless, when QTL analysis was performed, complementary QTL alleles could also be identified in the two species.

Significant overdominance/underdominance was observed for several QTL (7%). However, it was not possible to causally relate these QTL to either the occurrence of heterosis in the F_1 hybrids or appearance of transgressive individuals in the F_2 generation. When two of these overdominant QTL were transferred to an advanced backcross generation (BC2), they continued to exert the same genetic effect as seen in the F_2 population. The same gene action was confirmed for one of those overdominant QTL when transferred to a BC2F3 generation. Unlike with the F_2, in the BC2 and BC2F3 populations significant number of transgressive individuals did appear which were directly related to either two (BC2) or one (BC2F3) of the overdominant QTL. Presumably at that point, many other segregating QTL had been eliminated (compared with the F_2) and these two overdominant QTL then had large enough effects to give rise to individuals with extreme phenotypes. These results suggest that, while overdominance may not be a major factor contributing to wide-cross transgression (compared with the occurrence of complementary QTL), this genetic mechanism can lead to transgressive segregation. It should be also pointed out that, while it is clear that some QTL (such as *dw1* and *dw2*) have a significant overdominant/underdominant gene action, we cannot determine from these data whether this is due to single gene overdominance or pseudo-overdominance due to tight linkage of dominant and recessive alleles of opposite effects. Studies are currently underway to distinguish these possibilities via high resolution QTL mapping of *dw1* and *dw2*.

The inability to detect significant two-locus interactions between QTL and other loci in the F_2 generation suggests that epistasis was not a major factor contributing to transgressive segregation. This conclusion is supported by the fact that *dw1* and *dw2* in the BC2 and *dw1* also in the BC2F3 generation continued to exert the same effects as seen in the F_2 generation. If the effects of these two QTL had been dependent on the presence of *P* alleles elsewhere in the genome one might expect that the loss of such *P* alleles during the backcrossing process would have modified or eliminated the effects of *dw1* and *dw2*.

Evolutionary implications: Interspecific hybridization and subsequent introgression of genes from the genome of one species into that of another occurs in nature, and in the case of crop plants, takes place frequently at the hands of plant breeders (RIESEBURG and BRUNSFELD 1991; HARLAN 1976). The evolutionary significance of interspecific introgression has been the subject of much debate (ANDERSON 1949). Conservatively, introgression is thought to lead to new races or species with adaptations and characteristics *intermediate* between the parental species (STEBBINS 1950). At the other extreme, it has been proposed that interspecific hybridization can lead to new races or species with characteristics or adaptations that exceed those of the parents. BRAINERD (1924) demonstrated the transgressive generation of "new" plant types in interspecific crosses in the genus *Viola* and proposed the possible birth of a new species in the naturally occurring offspring of one such cross *Viola affinis* × *Viola sagittata*. Likewise, LEWONTIN and BIRCH (1966) provided evidence that introgressive hybridization in tephritid flies can lead to new biotypes with adaptation to extreme temperatures.

While the research described here involved artificial hybridization, it nonetheless demonstrates that distinct, but crossable, species are likely to contain complementary alleles for many if not most characters. Moreover, novel combinations of these alleles can readily produce phenotypes that greatly exceed the range of the parents. This work thus lends credibility to the proposal that interspecific hybridization can result in relatively rapid adaptations to new environments (LEWONTIN and BIRCH 1966). Left unanswered by this research is the frequency with which interspecific transgression has actually led to speciation in nature and the magnitude to which transgressive hybridization has played a role in the evolution of new adaptive functions.

Implications for the genetic improvement of crop plants: For most crop species , the largest reservoir of genetic variation exists not within modern cultivars, but in their feral counterparts. In tomato it has been estimated that less than 5% of the available genetic variation exists in cultivars and the remainder is found in wild relatives (MILLER and TANKSLEY 1990). Wild relatives have been exploited as a source of disease and insect resistance in many crops, but their use beyond this has been rather limited (HARLAN 1976). Breeders have been reluctant to use wild species for two main reasons. First, while it is often possible, using conventional breeding methods, to transfer gene(s) for desirable traits from wild species into cultivars, the transfer is often accompanied by undesirable, linked genes, a phenomenon referred to as linkage drag (ZEVEN, KNOTT and JOHNSON 1983). Even after many backcross generations and selection, selected genes are often accompanied by linked DNA segments large enough to carry hundreds of undesirable genes (YOUNG and TANKSLEY 1989). The use of molecular linkage maps in crops has provided a way to monitor and facilitate interspecific gene transfer and to mitigate linkage drag, improving the prospects for successful introgression of desirable genes from wild spe-

cies (TANKSLEY *et al.* 1989; YOUNG and TANKSLEY 1989).

A second difficulty with using wild germplasm is in identifying accessions that contain the trait of interest. Normally germplasm is screened for the trait of interest, and those accessions demonstrating the trait are then used as parents in backcrossing schemes to the cultivated species. While this approach is straightforward for characters such as disease resistance, it is often difficult to apply successfully for other characters. Yield, for example, is not a character than can be readily measured in wild germplasm. Wild species often have characteristics, selected against in modern cultivars (*e.g.*, prostrate growth habits or seed shattering), that cause them to yield much less than their cultivated counterparts. In tomato, for example, many of the wild species are self-incompatible and will not set fruit under commercial field situations where natural pollinating insects are absent. If one were to measure yield in these species, it would be concluded that they are not a potential source of yield genes.

The study reported here suggests that, while wild species may not display a character of interest (or have reduced amounts of that character compared with their cultivated counterparts), it is likely that they possess alleles that can improve the character. For all of the characters examined in this study, at least one QTL was identified in the wild species that had an allellic effect opposite of what would be expected based on the phenotype of the wild plant. Moreover, two overdominant QTL, *dw1* and *dw2*, were identified, which when backcrossed into the cultivated tomato, positively affected characters of agronomic importance (growth rate and fruit yield). This result was confirmed again for *dw1* in a more advanced generation. Reports of transgressive variation for yield from wild species suggests that this is not an isolated phenomena (FREY *et al.* 1981). By combining wild species crosses with molecular linkage maps, it should be possible to identify and selectively introgress new alleles that can improve characters for which wild species are normally not considered to be a source of useful variation. This may be especially important in self-pollinated crops like tomato, cotton, soybean and wheat where the existing cultivated germplasm base for breeding is narrow and would likely benefit by the infusion of new alleles.

This work was supported in part by grants from the National Research Initiative Competitive Grants Program, U.S. Department of Agriculture No. 91-37300-6418 and by the Binational Agricultural Research and Development Fund (No. IS-1822-90C). Thanks to G. CHURCHILL, S. MCCOUCH, A. PATERSON, M. SORRELLS and D. VIANDS for helpful comments, MICHAEL BRIGGS and ELIANA YGLESIAS for technical assistance and to STEFFIE DAVID for help in preparing the manuscript.

LITERATURE CITED

ANDERSON, E., 1949 *Introgressive hybridization*. John Wiley & Sons, New York.

BERNATZKY, R., and S. D. TANKSLEY, 1986 Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genetics **112:** 887–898.

BRAINERD, E., 1924 Some natural violet hybrids of North America. Vermont Agr. Exp. Sta. Bull. No. **239**.

DARLINGTON, C. D., and K. MATHER, 1949 *The Elements of Genetics*. Allen & Unwin, London.

DE VICENTE, M. C., and S. D. TANKSLEY, 1991 Genome-wide reduction in recombination of backcross progeny derived from male *versus* female gametes in an interspecific cross of tomato. Theor. Appl. Genet. **83:** 173–178.

EDWARDS, M. D., C. W. STUBER and J. F. WENDEL, 1987 Molecular-marker-facilitated investigations of quantitative-trait loci in maize. I. Numbers, genomic distribution, and types of gene action. Genetics **116:** 113–125.

ENGELS, W. R., 1983 The P family of transposable elements in Drosophila. Annu. Rev. Genet. **17:** 315–344.

FEINBURG, A. P., and B. VOGELSTEIN, 1983 A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132:** 6–13.

FREY, K. J., T . S. COX, D. M. RODGERS and P. BRAMEL-COX, 1981 Increasing cereal yields with genes from wild and weedy species. Journal Paper No. J-11254 of the Iowa Agriculture and Home Economics Experimental Station, Ames, IA 50011. Project 2447, p. 51–68.

HARLAN, J. R., 1976 Genetic resources in wild relatives of crops. Crop Sci. **16:** 329–333.

KOSAMBI, D. D., 1944 The estimation of map distances from recombination values. Ann. Eugen. **12:** 172–175.

LANDER, E. S., and D. BOTSTEIN, 1989 Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics **121:** 185–199.

LANDER, E. S., P. GREEN, J. ABRAHAMSON, A. BARLOW, M. J. DALY, S. E. LINCOLN and L. NEWBURG, 1987 Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1:** 174–181.

LEWONTIN, R. C., and L. C. BIRCH, 1966 Hybridization as a source of variation for adaptation to new environments. Evolution **20:** 315–336.

LOTSY, J. P., 1916 *Evolution by Means of Hybridization*. M. Nijhoff, The Hague.

MILLER, J. C., and S. D. TANKSLEY, 1990 RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. Theor. Appl. Genet. **80:** 437–448.

PATERSON, A. H., S. DAMON, J. D. HEWITT, D. ZAMIR, H. D. RABINOWITCH, S. E. LINCOLN, E. S. LANDER and S. D. TANKSLEY, 1991 Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. Genetics **127:** 181–197.

REEVES, R. G., and A. J. BOCKHOLT, 1964 Modification and improvement of a maize inbred by crossing it with Tripsacum. Crop Sci. **4:** 7–10.

RICK, C. M., and P. G. SMITH, 1953 Novel variation in tomato species hybrids. Am. Nat. **88:** 359–373.

RIESEBERG, L. H., and S. J. BRUNSFELD, 1992 Molecular evidence and plant introgression, pp. 151–176 in *Molecular Sysematics in Plants*, edited by P. E. SOLTIS, P. S. SOLTIS and J. J. DOYLE. Chapman & Hall, New York.

SAS Institute, Inc., 1988 *SAS Users Guide: Statistics*. SAS Institute, Cary, N.C.

STEBBINS, G. L., JR., 1950 *Variation and Evolution in Plants*. Columbia University Press, New York.

STUBER, C. W., M. D. EDWARDS and J. F. WENDEL,

1987 Molecular marker-facilitated investigations of quantitative trait loci in maize. II. factors influencing yield and its component traits. Crop Sci. **27:** 639–648.

TANKSLEY, S. D., N. D. YOUNG, A. H. PATERSON and M. W. BONIERBALE, 1989 RFLP mapping in plant breeding: new tools for an old science. BioTechnology **7:** 257–264

TANKSLEY, S. D., M. W. GANAL, J. P. PRINCE, M. C. DEVICENTE, M. W. BONIERBALE, P. BROUN, T. M. FULTON, J. J. GIOVANONNI, S. GRANDILLO, G. B. MARTIN, R. MESSEGUER, J. C. MILLER, L. MILLER, A. H. PATERSON, O. PINEDA, M. RODER, R. A. WING, W. WU and N. D. YOUNG, 1992 High density molecular linkage maps of the tomato and potato genomes: biological inferences and practical applications. Genetics **132:** 1141–1160.

VEGA, U., and K. J. FREY, 1980 Transgressive segregation in inter and intraspecific crosses of barley. Euphytica **29:** 585–594.

YOUNG, N. D., and S. D. TANKSLEY, 1989 Restriction fragment length polymorphism maps and the concept of graphical genotypes. Theor. Appl. Genet. **77:** 95–101.

ZAMIR, D., and Y. TADMOR, 1986 Unequal segregation of nuclear genes in plants. Bot. Gaz. **147:** 355–358.

ZEVEN, A. C., D. R. KNOTT and R. JOHNSON, 1983 Investigation of linkage drag in near isogenic lines of wheat by testing for seedling reaction to races of stem rust, leaf rust and yellow rust. Euphytica **32:** 319–327.

Communicating editor: D. CHARLESWORTH