Final report for Grant DE-FG02-91ER20037

DOE/ER/20037--71

4/30/97

Martha A. Mutschler
Dept. of Plant Breeding
Cornell University
Ithaca, N.Y.

The work supported by this project is divided into studies of the interspecific barriers hybrid breakdown (see topic 1 below), segregation distortion (see topic 2 below) and unilateral incongruity (see topics 3 and 4 below).

Topic 1. Reproductive abnormalities reduced the percent stainable pollen, and fruit and seed set in interspecific F_2 populations derived from crosses of *Lycopersicon esculentum* and *L. pennellii* but were not observed in parental lines and interspecific F_1's. The degree to which these reproductive abnormalities were expressed in the interspecific F_2 populations was affected by cytoplasm. Reproduction was impeded in interspecific F_2 populations containing *L. esculentum* cytoplasm (F_2^{Le}) by reduction in pollen production, the lack of fruit set and a high proportion of parthenocarpic fruit among plants capable of fruit set. The F_2 populations containing *L. pennellii* cytoplasm (F_2^{Lp4}) showed a reduced frequency reproductive abnormalities at all stages of reproductive development, resulting in higher values for percent stainable pollen, and fruit and seed set and in higher proportions of the populations being capable of setting fruit or seed in the F_2^{Lp4} than in F_2^{Le} population. The major barrier remaining in F_2^{Lp4} populations was reduced fruit set compared to parental lines. The barrier to fruit and seed set observed in the F_2^{Le} populations, and to a lesser extent in the F_2^{Lp4} populations, occurs around the time of fertilization or early embryonic development. The effect of *L. pennellii* cytoplasm on barriers in the F_2^{Lp4} populations could result from interaction between cytoplasm and nuclear genes during fertilization of the F_1 plants to produce F_2 populations, and/or during the reproductive development, self-pollination, and/or fertilization of the F_2 and subsequent generation. This work resulted in the publication Liu et al 1995 Alterations of the manifestations of hybrid breakdown in *Lycopersicon esculentum* x *L. pennellii* F_2 populations containing *L. esculentum* versus *L. pennellii* cytoplasm Sexual Plant Reproduction 8:361-368. A reprint of this paper is appended

Topic 2. Deficiency of *Lycopersicon esculentum* allele (E) was observed from the RFLP and isozyme data of the F_2 populations derived from the cross *L. esculentum* x *L. pennellii*. The genome composition of the F_2 populations containing *L. pennellii* cytoplasm (F_2^{Lp4}) has a lower proportion of the homozygous *L. pennellii* (PP) genotypes and a higher proportion of heterozygote (EP) genotypes than that of the F_2 populations containing *L. esculentum* cytoplasm (F_2^{Le}). A lower proportion of the *L. pennellii* alleles (P) was also observed in F_2^{Lp4} as compared to F_2^{Le} when each marker locus was tested individually. To study the effects of gametic and zygotic

Topic 4 Genetic studies were also performed on the control of interspecific reproductive barriers, such as unilateral incongruity (UI). Male fertile plants were chosen from two F_2 populations derived from a cross between *L. esculentum* or an isocytoplasmic *L. esculentum* plant and *L. pennellii*. Pollen tube growth was classified as congruous or incongruous in crosses between these F_2 plants and the two species to classify the male and female expression of UI in the F_2 plants. Congruous pollen tube growth was found for some F_2 plants in crosses which generally are incongruous suggesting autosomal control of this trait. There was no effect of cytoplasm on UI. Segregation for molecular markers was compared with male and female expression of UI; two and eight markers were associated with UI, respectively, after removing markers with skewed segregation ratios. Markers on chromosomes 4 and 11 were associated with UI in both the staminate and pistillate parent. These results are discussed in relation to prior work on crossing barriers. Due to overlapping male and female functions in F_2 populations, future work should use backcross populations to assist in dissecting the male and female components of UI. A manuscript containing this work and additional data is in preparation.

Summary. Deficiency of *Lycopersicon esculentum* allele (E) was observed from the RFLP and isozyme data of the F_2 populations derived from the cross *L. esculentum* x *L. pennellii*. The genome composition of the F_2 populations containing *L. pennellii* cytoplasm(F_2^{Lp4}) has less abundant proportion of homozygous *L. pennellii* (PP) genotype and a higher proportion of heterozygote (EP) genotype than that of the F_2 populations containing *L. esculentum* cytoplasm (F_2^{Le}). A decrease in favoring *L. pennellii* allele (P) was also observed in F_2^{Lp4} when each marker locus was tested by the likelihood-ratio chi-squared tests (G^2). To study the association of segregation distortion with gametic and zygotic selection, the patterns of significance expected for the G^2 tests at the marker loci were derived for ten selection models specifying gametic or zygotic selection at a hidden locus linked to these marker loci. Segregation distortion caused by four of the selection models studied can be uniquely identified by the patterns of significance expected for the G^2 tests at the marker loci. Comparison of the chromosomal regions associated with specific selection models across populations (of this experiment and previous publications) indicated that the segregation distortion observed in chromosome 10 is associated with a zygotic selection model affecting both arms of the chromosome, and cytoplasm substitution could have an effect on decreasing the segregation distortion on the long arm of the chromosome.

Key words: *Lycopersicon*-Genome composition-Segregation distortion-Gametic selection-Zygotic selection

Introduction

Aberrant segregation ratios have been reported in the progeny of numerous interspecific hybrids including those from the genera *Gossypium*, *Lycopersicon*, *Lens*, *Capsicum*, *Cucurbita*, *Solanum*, and *Oryza* (Stephens 1950; Rick 1963; Rick 1969; Rick 1972; Weeden and Robinson 1986; Zamir and Tadmor 1986; Bonierbale et al. 1988; Sato et al. 1990). Zamir and Tadmore (1986) estimated that the percentage of loci exhibiting segregation distortion for interspecific crosses was approximately fourfold (54%) that for intraspecific crosses (13%) in the genera *Lens*, *Capsicum*, and *Lycopersicon*. Most cases of segregation distortion observed in the progeny of interspecific hybrids were thought to result from genes subjected to directional selection at either pre- or post-zygotic stages of plant development rather than abnormal chromosome pairing during meiosis in the F_1 hybrids (Rick 1969; Weeden and Robinson 1986; Zamir and Tadmor 1986).

Segregation distortion in F_2 populations or other progeny derived from *L. esculentum* x *L. pennellii* crosses was reported for morphological single gene traits and molecular markers (Rick 1969; Rick 1972; Bernatzky and Tanksley 1986; Zamir and Tadmor 1986; Gadish and Zamir 1987; Kinzer et al. 1990; Chetelat and DeVerna 1991; de Vicente 1991). Most of these populations studied exhibited a deficiency of the *L. esculentum* allele in the distorted segregation ratios. Only one backcross population generated by pollinating an F_1 plant with *L. pennellii* pollen had equal numbers of loci showing segregation distortion favoring *L. esculentum* allele and *L. pennellii* allele (de Vicente 1991). The deficiency of *L. esculentum* alleles combined with the reduction of recombination in the male gametes of the hybrid (Rick 1969; Rick 1972; de Vicente and Tanksley 1991) would result in the tendency to preserve and favor large linkage blocks of *L. pennellii* chromosomes in the progeny and impede the introgression of small linkage blocks into tomato (Rick 1969). Systematic studies of the fundamental nature of segregation distortion are necessary to avoid the influences of segregation distortion in breeding programs involving interspecific crosses and the introgression of genes from wild relatives to crop species.

Segregation distortion was frequently observed in interspecific F_2 populations (Stephens 1950; Rick 1963; Rick 1969; Rick 1972; Weeden and Robinson 1986; Zamir and Tadmor 1986; Bonierbale et al. 1988; Sato et al. 1990), but these studies were limited to reporting the deviating segregation ratios. Wagner et al. (1992) also estimated linkage between marker loci showing distorted segregation and loci under the control of selection, but only limited cases of complete selection in gametic and zygotic stages were studied. In this study, we attempt to determine the mechanism(s) responsible for segregation distortion, to identify the systematic deviations for segregation ratios resulting from selection forces at different developmental stages, and to examine the possibility of cytoplasmic influence on segregation distortion.

Materials and Methods

Plant materials

Two types of F_2 populations were used. One type of F_2 population, designated "F_2^{Le}", was generated from F_1^{Le}, an *Lycopersicon esculentum* x *Lycopersicon pennellii* F_1 of which the *L. esculentum* parent is the cultivar New Yorker and the *L. pennellii* parent is the self compatible accession of *L. pennellii* 'LA716' (Lp) (PI 246502, from Dr. C. M. Rick, University of California, Davis). The other type of F_2 population, designated "F_2^{Lp4}", was generated from F_1^{Lp4}, an *L. esculentum* x *L. pennellii* F_1, of which the *L. pennellii* parent was still LA716 but the *L. esculentum* parent was New Yorker^{Lp4} (NY^{Lp4}), a plant derived from the *L. esculentum* line New Yorker but containing the cytoplasm of *L. pennellii* LA716 (Mutschler 1990). Most accessions of *L. pennellii* are self incompatible, but since LA716 possesses a dominant form of self compatibility, the F_1^{Le} and F_1^{Lp4} plants derived from it are also self compatible (Liu et al. in press), reducing or eliminating the effect of the *S* locus on the segregation ratios of its linked loci in the F_2 populations. Populations III (an F_2^{Le}) and IV (an F_2^{Lp4}), with 53 and 60 plants respectively, were grown simultaneously in a greenhouse (Liu et al. in press). Population VI (an

F_2^{Lp4}) with 144 plants, was planted in the same greenhouse in the subsequent year (Mutschler et al. submitted).

Linkage analysis of isozyme and RFLP markers

Leaf extracts of plants in populations III and IV were assayed electrophoretically for monogenic segregation of genes *Tpi-2*, *Pgi-1*, and *Pgm-2* coding for the enzymes, triosephosphate isomerase (TPI, E.C. 5.3.1.1), phosphoglucoisomerase (PGI, E.C. 5.3.1.9), and phosphoglucomutase (PGM, E.C. 2.7.5.1), respectively, as described in Stephenson (1991). The agarose stain mixture for TPI contained 4 ml of 0.1 M Tris, pH 8.0, 3 mg of NAD, 80 mg sodium arsenate, 8 mg EDTA, 2 mg dihydroxyacetone phosphate, 40 units glyceraldehyde 3-phosphate dehydrogenase, 3 mg MTT, 0.2 mg PMS, and 5 ml of 1.6% agarose.

Genomic DNA was extracted from plants using the method of Doyle and Dickson (1987). RFLP analysis was performed as in Mutschler et al. (submitted) using genomic clones of known chromosomal locations (Dr. S. D. Tanksley) chosen to cover the entire genome. Ninety-seven RFLP markers were analyzed for populations III and IV. One hundred eighty-seven RFLP markers were analyzed for population VI, including 82 of the markers used in populations III and IV.

Linkage maps of populations III, IV, and VI were derived using computer software MAPMAKER/EXP v.3.0 (Lincoln and Lander 1992; Lincoln et al. 1993) and have average space between markers of 14.4 cM, 15.8 cM, and 6.6 cM, respectively.

Genome composition

The *L. esculentum* genome content, i.e. the percentage of the total genome comprised of *L. esculentum* alleles, was estimated for each plant in the three F_2 populations using computer program HYPERGENE v.0.9 (Young and Tanksley 1989) as described in Mutschler et al. (submitted). The genome percentage of the EE, EP, and PP genotypes, i.e. the percentages of

total genome that were homozygous for *L. esculentum*, heterozygous for *L. esculentum* and *L. pennellii*, and homozygous for *L. pennellii*, respectively, were similarly estimated for each plant in all three F_2 populations.

G^2 goodness-of-fit test

G^2 goodness-of-fit test (Agresti 1990) is the likelihood-ratio chi-squared test used for testing monogenic segregation (1:2:1 for EE: EP: PP) of each isozyme and RFLP marker. This df=2 G^2 test was partitioned to two df=1 G^2 tests: the 1:1 G^2 test for allele frequencies and the e^2:2ep:p^2 G^2 test for the predicted zygotic frequencies. The 1:1 G^2 test is a goodness-of-fit test for single locus allele segregation (1:1 for *L. esculentum* allele (E) and *L. pennellii* allele (P)). Significant deviations from the expected 1:1 allele frequencies indicate unequal frequencies for E allele and P allele at the marker loci in the F_2 population. The e^2:2ep:p^2 G^2 test is a goodness-of-fit test for random fertilization under the assumption that allele frequencies for E allele and P allele in the male and female gametes are the same as those in the F_2 population, resulting in a predicted segregation ratio of e^2:2ep:p^2 for EE: EP: PP, where e and p are the frequencies of E allele and P allele, respectively, at the locus in the F_2 population. The patterns of the significance of these G^2 statistics were used to investigate the segregation distortion observed in a group of linked markers (discussed below).

Results and Discussion

Genome composition

Most of the markers used showed codominant segregation for EE, EP, and PP genotypes in all three F_2 populations. However, eight of the RFLP markers showing codominant segregation in population III did not segregate in populations IV and VI. The published tomato linkage map indicates that these eight markers are all located on chromosome 1 between CT233 and TG71 (Tanksley 1992). The homogeneity of this chromosomal region for *L. pennellii* alleles in the F_2^{Lp4} populations is explained by the fact that the NY^{Lp4} and F_1^{Lp4} progenitors of the F_2^{Lp4}

populations had EP and PP genotypes, respectively, for all of the markers in this region (Mutschler et al. submitted). The non-segregation of this region from the F_2^{Lp4} populations should not affect the analyses of segregation distortion at loci in the rest of the genome, but it was excluded from consideration in all three populations for equity in the comparisons of genome composition among populations.

The genomes of the plants in populations III, IV and VI had mean *L. esculentum* contents of 45.1%, 46.8%, and 46.3%, respectively (Figure 1). The mean *L. esculentum* content of each of the three populations was significantly smaller than the expectation of 50% (population III: $t_{df=52} = -5.66$, $p < 0.001$; population IV: $t_{df=59} = -3.49$, $p < 0.001$; population VI: $t_{df=143} = -6.74$, $p < 0.001$). However, there was no statistical significance for the difference in the mean *L. esculentum* content among the three populations ($F^{df1=2}_{df2=254} = 1.03$, $p = 0.36$).

The plants in population III had mean genome percentages of the EE, EP, and PP genotypes of 19.2%, 51.7%, and 29.1% (Figure 2), compared with the expectation of 25% ($t_{df=52} = -5.96$, one-tailed $p < 0.001$), 50% ($t_{df=52} = 1.27$, two-tailed $p = 0.27$), and 25% ($t_{df=52} = 3.40$, one-tailed $p < 0.001$), respectively. The genome percentage of EP in population III meets expectation, but the observed genome percentages of EE and PP are significantly lower and higher, respectively, than expected. The plants in population IV had genome percentages of the EE, EP, and PP genotypes of 20.1%, 53.4%, and 26.5% (Figure 2), compared with the expectation of 25% ($t_{df=59} = -4.94$, one-tailed $p < 0.001$), 50% ($t_{df=59} = 3.04$, one-tailed $p = 0.001$), and 25% ($t_{df=59} = 1.31$, two-tailed $p = 0.20$), respectively. The genome percentage of PP in population IV meets expectation, but the observed genome percentages of EE and EP are significantly higher than expected. In population IV the mean genome percentage of PP was significantly higher than the mean genome percentage of EE ($t_{df=116} = 4.22$, one-tailed $p < 0.001$). The plants in population VI had genome percentages of the EE, EP, and PP genotypes of 19.4%, 53.8%, and 26.8% (Figure 2), all of which are significantly different from the expectations of 25% ($t_{df=143} = -9.36$, one-tailed $p < 0.001$), 50% ($t_{df=143} = 5.34$, one-tailed $p < 0.001$), and 25% ($t_{df=143} = 2.59$, one-tailed p =

0.005), respectively. Therefore, both F_2^{Lp4} populations (IV and VI) had significantly higher heterozygosity, and all three populations had lower mean genome percentage of EE and higher mean genome percentage of PP, resulting in all of their mean *L. esculentum* contents being significantly lower than 50%. However, since the genome percentages of the EE, EP, and PP genotypes sum up to one for each plant (Figure 3), the three *t* tests for these three genome percentages in each population are not independent.

A triangle diagram illustrates the deviation of the genome composition from the expectation by locating each plant in the triangle with its genome percentages of the EE, EP, and PP genotypes used as the coordinates .(Figure 3). A simulated population of 150 plants was 'generated for comparison with the genome percentages of the EE, EP, and PP genotypes of populations III, IV, and VI. The simulated population was derived such that the genome percentages of the EE, EP, and PP genotypes of its individual were randomly generated from normal distributions with standard errors of 8.0 (the mean standard error of the percentage genome of EE, EP, and PP genotypes of population VI) and means of 25%, 50%, and 25%, respectively. All three experimental populations showed a striking deficiency of EE genotype compared to the simulated population.

For each of the EE, EP, and PP genotypes, the mean genome percentages of populations III, IV, and VI were compared by an *F* test, within which two orthogonal contrasts (L_1 and L_2) of the means were planned: L_1 for comparing the mean genome percentage of population III with those of populations IV and VI; L_2 for comparing the mean genome percentage of populations IV with that of population VI. The three *F* tests for the EE, EP, and PP genotypes were not independent, since the genome percentages of the EE, EP, and PP genotypes sum up to one for each plant. The tests derived $F^{df\,1=2}_{df\,2=254} = 0.22$, p = 0.80 for the mean genome percentages of the EE genotype with $t_{df=254} = -0.48$, one-tailed p = 0.32 for L_1, and $t_{df=254} = 0.62$, one-tailed p = 0.27 for L_2; $F^{df\,1=2}_{df\,2=254} = 1.07$, p = 0.34 for the mean genome percentages of the EP genotype with $t_{df=254} = -1.38$, one-tailed p = 0.08 for L_1, and $t_{df=254} = -0.23$, one-tailed p = 0.41 for L_2; and

$F^{df\,1=2}_{df\,2=254}$ = 1.62, p = 0.20 for the mean genome percentages of the PP genotype with $t_{df=254}$ = 1.82, one-tailed p = 0.03 for L_1, and $t_{df=254}$ = -0.23, one-tailed p = 0.41 for L_2. The genome compositions of populations IV and VI are more similar to each other than to population III in having lower mean genome percentages of PP genotype, which contribute an increase in the mean genome percentages of EP genotype in populations IV and VI. The similarity in genome composition between populations IV and VI indicates that the $F_2{}^{Lp4}$ populations had less abundant proportions of PP genotype (and higher proportions of EP genotype), which indicates a reduce in the deficiency of E allele in population IV and VI compared to that in population III.

Segregation distortion

Twenty marker loci in 9 chromosomal regions, 4 marker loci in 2 chromosomal regions, and 58 marker loci in 11 chromosomal regions showed significant deviations from the expected 1:1 allele frequencies in population III (Table 1), population IV (Table 1), and population VI (Figure 1), respectively. All the significant loci in populations III and IV, and 54 of the significant loci in 9 chromosomal regions of population VI showed significant deviations with an overabundance of the P allele. The remaining 4 marker loci in 2 chromosomal regions of population VI favored the E allele.

In population III, 16 markers showed segregation distortion for the expected 1:2:1 genotypic frequencies (Table 1). Twelve of these markers, in 7 chromosomal regions, showed significant skew for the 1:1 G^2 test but not for the e^2:2ep:p^2 G^2 test. In each of these cases, the skew favored the P allele. Eight markers in population IV showed segregation distortion for the expected 1:2:1 genotypic frequencies (Table 1). Three of these markers, in one chromosomal region, showed significant skew for the 1:1 G^2 test but not for the e^2:2ep:p^2 G^2 test, all of which favored the P allele (Table 1). The observations in populations III and IV of aberrancies in the 1:1 allele segregation ratios favoring the P allele agree with prior reports of aberrant segregation for molecular markers and morphological single gene traits in *L. esculentum* x *L. pennellii* F_2

populations (Zamir and Tadmor 1986; Kinzer 1989; Chetelat and DeVerna 1991; de Vicente 1991) and other populations derived from crosses between *L. esculentum* and *L. pennellii* (Rick 1969; Rick 1972). However, population IV exhibited less serious aberrancies in the 1:1 allele frequencies than population III, showing fewer marker loci with deviated segregation ratios and weaker deviation in segregation ratios for the marker loci which had skewed allele frequencies in both populations. This indicates that cytoplasm substitution could affect the mechanism controlling segregation distortion, decreasing the extent to which the P allele is favored (as observed in the comparison of genome composition). Seventy-one marker loci in population VI showed segregation distortion for the expected 1:2:1 genotypic frequencies, thirty-four of these marker in 8 chromosomal regions showed significant skew for the 1:1 G^2 test but not for the e^2:2ep:p^2 G^2 test, all of which favored the P allele (Figure 4). It was expected that population VI would have more chromosomal regions with markers showing significant deviation in segregation than populations III and IV, because a larger population is statistical more powerful in detecting deviations from the expected Mendelian monogenic segregation ratios. The chromosomal regions identified in population VI by marker loci with significantly deviated segregation ratios included all the significant chromosomal regions identified in population IV, and five of the seven chromosomal regions identified in population III (Table 1 and Figure 4).

Association of the patterns of the significance of G^2 test results and the nature of segregation distortion

The possible causal factors of segregation distortion were examined for association with the patterns of significance of the 1:2:1, 1:1, and e^2:2ep:p^2 G^2 tests. A locus *A* was designated as the origin of segregation distortion, and a locus *B*, which is linked to the *A* locus by a recombination fraction *r*, was designated as a marker locus showing distorted segregation ratios under the effect of the *A* locus. If selections at the *A* locus occurs in the gametic stage with a proportion *p* of the

"A" genotype female gametes and a proportion q of the "A" genotype male gametes being eliminated, the progeny will have an expected genotype ratio of

$$\mathrm{AA:Aa:aa} = (1-p)(1-q):2(1-pq):(1+p)(1+q),$$

where $-1 < p, q < 1$. Instances in which $-1 < p, q < 0$ indicate that proportions p and q of the "a" genotype female and male gametes, respectively, are eliminated. Since B locus is linked to A locus by r recombination fraction, the expected genotype ratio of the progeny was derived as

$$\mathrm{BB:Bb:bb} = \left[1-p(1-2r)\right]\left[1-q(1-2r)\right]:2\left[1-pq(1-2r)^2\right]:\left[1+p(1-2r)\right]\left[1+q(1-2r)\right].$$

The segregation ratios at B locus and a known population size n enable us to derive the expected results of the 1:2:1, 1:1, and e^2:2ep:p^2 G^2 tests at B locus for each A locus with segregation ratios controlled by a hypothesized selection model in the gametic stage (specified by p and q). Similarly, if selections at the A locus occur in the zygotic stage with a proportion of x of the "AA", a proportion of y of the "Aa", and a proportion of z of the "aa" genotypes being eliminated, the progeny will have a genotype ratio of

$$\mathrm{AA:Aa:aa} = (1-x):2(1-y):(1-z),$$

where $0 < x, y, z < 1$. A B locus linked with the A locus by a recombination fraction r will have a progeny genotype ratio of

$$\begin{aligned}\mathrm{BB:Bb:bb} = &\left[(1-x)(1-r)^2+(1-y)2r(1-r)+(1-z)r^2\right]\\ &:2\left[(2-x-z)r(1-r)+(1-y)\left(r^2+(1-r)^2\right)\right]\\ &:(1-x)r^2+(1-y)2r(1-r)+(1-z)(1-r)^2.\end{aligned}$$

Therefore, the expected results of the 1:2:1, 1:1, and e^2:2ep:p^2 G^2 tests at B locus can be derived from the segregation ratios at B locus and a known population size n for each A locus with

segregation ratios controlled by a hypothesized selection model in the zygotic stage (specified by x, y, and z).

The results of the G^2 tests at a B locus (i.e. the marker locus) in each hypothesized model are affected by a combination of factors including population size, distance from A locus (i.e. locus under selection), and the magnitude of the selection force at the linked A locus. A larger population should be expected to show higher significance levels, because the larger sample size is more powerful statistically. Also, a smaller distance between an A locus and a B locus and/or a stronger selection force at an A locus will result in higher significance levels for the results of the G^2 tests at skewed B loci due to the effect of linkage. Therefore, we expect to observe the most significant G^2 tests results for B loci closest to an A locus, with decreasing significance of the G^2 tests results at B loci farther away from the A locus. By varying population size, distance from an A locus, and the magnitude of the selection force at an A locus, we derived the patterns of significance of the G^2 tests results and the expected frequency of heterozygote for hypothesized models of selection in gametic stage (Table 2) and zygotic stage (Table 3). The selection models studied are

A: gametic selection against the same genotype in both sexes,

B: gametic selection against different genotypes in the two sexes with similar proportions of gametes being eliminated,

B': gametic selection against different genotypes in the two sexes with the proportion of gametes being eliminated in one sex being less than two times of that in the other sex,

B": gametic selection against different genotypes in the two sexes with the proportion of gametes being eliminated in one sex being more than two times of that in the other sex,

C: gametic selection against one genotype in only one of the two sexes,

D: dominant zygotic selection against only the "AA" and "Aa" genotypes,

E: recessive zygotic selection against only the "aa" genotype,

F: overdominant zygotic selection against only the "Aa" genotype,

G: underdominant zygotic selection against only the "AA" and "aa" genotypes,

and

H: codominant zygotic selection against only the "AA" and "Aa" genotypes.

It is not possible to uniquely identify the mechanism of segregation distortion from segregation ratio data of the type available here. However, some mechanisms of segregation distortion can be identified by the patterns of significance expected for the G^2 tests. Segregation distortion caused by selection models A, B', D, and F can be uniquely identified by the patterns of significance expected for the G^2 tests (Table 2 and 3). Some of the other selection models derive similar patterns of significance due to sharing a specific relative magnitude of the eliminating factors that specify these models. For example, selection models B" and C share similar patterns of significance, because model B" is the intermediate model between model B and model C and approaches model C in the relative magnitude of their eliminating factors (p and q). Similar patterns of significance derived from different models can also result from their similar segregation ratios. Segregation ratios resulting from model C in gametic selection will resemble those generated from a special case of model H in zygotic selection, i.e. x=2y. Likewise, distorted segregation ratios with a much lower ratio for heterozygote genotype than the expected ratios can be generated by models B and G.

Analysis of the segregation distortion observed in populations III, IV and VI

Some variation in the observed genotype ratios is expected due to sampling, since a large number of tests (one at each marker) are performed. Furthermore, the individual tests are not independent for tightly linked markers. Therefore, an concordance of significant G^2 tests across a sequence of linked loci (*B* loci) is necessary to determine whether segregation distortion is caused

by linkage to a hidden locus (*A* locus) under selection. For a genetic linkage map with an average distance between markers of approximate 20 cM, the linkage effect of the selection at an *A* locus must extend to a minimum of 20 cM, i.e. covering a 40 cM interval, for at least two adjacent markers to show segregation distortion. A relatively strong selection force would be needed to give evidence of single locus selection in adjacent markers for a population of 50 individuals, e.g. populations III and IV. This force must cause a minimum elimination of 35% of the gametes from both sexes in model A, 85% of the gametes from one sex in models B, B', and B", 70% of the gametes from one sex in model C, 70% of the zygotes of one genotype in models D and H, or 100% of the zygotes of one genotype in models E, F, and G to result in significance. Therefore, populations III and IV can only be expected to detect selection in model A and in the most extreme cases of the other single locus selections. This explains why many isolated cases of segregation distortion were detected in populations III and IV, and why neither population is suitable for further analysis of segregation distortion based on the pattern of significance of the G^2 test. In contrast, a larger population of 140 individuals with an average distance between markers of approximately 20 cM can detect single locus selection in adjacent markers for moderate selection force. The force needs to result in a minimum elimination of 20% of the gametes from both sexes in model A, 70% of the gametes from one sex in models B, B', and B", 45% of the gametes from one sex in model C, 50% of the zygotes of one genotype in models D and H, or 70% of the zygotes of one genotype in models E, F, and G for significance to be detected. A higher density of marker loci also provides a better chance of demonstrating the complete pattern of significance for distorted segregation ratios. Therefore, population VI (with 144 plants and higher density of marker loci) was used for further analysis of segregation distortion based on the patterns of significance of the G^2 test.

Nine chromosomal regions with closely linked markers show significant segregation distortion in population VI (Figure 4). The region TG83-TG197 on chromosome 1 shows the pattern of significance for model B and model G. The *A* locus inferred for this region is likely to exert a weak selection force, since its effect extends to only 8.6 cM. The region TG165-TG502 on

chromosome 2 showed a mixture of different patterns of significance for the models B', B", E, and H. If the markers placed on the linkage map with lower confidence (LOD <3) were not considered, the patterns of significance in this chromosomal region seem to vary from model H near CT255 to model E near TG426. The fact that many markers were placed in this chromosomal region with lower confidence may suggest the effect of missing data, which can cause variation in the patterns of significance. However, the variation in the patterns of significance may also suggest the existence of more than one locus under selection in this chromosomal region, and/or that selection occurs at more than one developmental stage. The region CT22-TG152 on chromosome 3 showed patterns of significance for model B" and model E. The region TG623-CD64 on chromosome 5 showed the pattern of significance of model B'. The region TG25-TG548 on chromosome 6 showed the pattern of significance of model A. The region TG176-TG481 on chromosome 8 showed the pattern of significance for model B and model G. However, model B' could also be involved, since there is a 13.3 cM gap between the marker loci showing the highest significance. The region TG254-TG35 on chromosome 9 showed a mixture of patterns of significance for models A, D, and H. The 22.3 cM gap next to the most significant marker locus (TG223A) made it difficult to classify the pattern of significance. The region TG230-TG63 on chromosome 10 showed the pattern of significance of model H. The region TG344-TG400 on chromosome 11 showed the pattern of significance of model B'.

Four chromosomal regions, located on chromosomes 5, 6, 10, and 11, were identified in population VI as being associated with a unique selection model (models B', A, H, B', respectively). The significance patterns of these chromosomal regions in population VI were compared with those in populations III and IV. We attempted to compare selection models associated with specific chromosomal region across populations despite that fact that the patterns of significance in populations III and IV can not be determined as being generated exclusively by one selection model, since most significant regions identified in populations III and IV did not show complete patterns of significance. The chromosome 5 region identified in population VI also showed significant skewing in population III, which was also associated with the same selection

model. But this region was not significantly skewed in population IV. The chromosome 6 region identified in population VI also showed significant skewing in both populations III and IV, which was also associated with the same selection model. The chromosome 11 region identified in population VI did not show significant skewing in either population III or IV. The chromosome 10 region identified in population VI was unique in that it was also identified in both populations III and IV, and is associated with the same selection model (H) in all three populations. Similar patterns of significance that was associated with the zygotic selection model H can be obtained by analyzing the segregation ratios previously reported for other chromosome 10 marker loci in *L. esculentum* x *L. pennellii* F_2 populations (Kinzer et al. 1990; Chetelat and DeVerna 1991; de Vicente 1991) (Table 4). Segregation distortion was observed for the entire length of chromosome 10 in population III. Since the highest significance appears at the distal end of the short arm of chromosome 10 in population III, and linkage effect is not expected to extend to the other end of the chromosome, it is likely that another locus is under selection on the long arm of chromosome 10 in this population. The patterns of significance in the data from of Kinzer et al. (1990) and de Vicente (1991) also support the existence of two loci under selection on chromosome 10 (Table 4). In contrast, population IV, which has the *L. pennellii* cytoplasm instead of *L. esculentum* cytoplasm as population III and the F_2 populations of Kinzer et al. (1990) and de Vicente (1991), showed significant segregation distortion only on the distal end of the short arm of chromosome 10. This suggests that cytoplasm substitution could have an effect on decreasing the segregation distortion on the long arm of chromosome 10. However, population VI, which also has the *L. pennellii* cytoplasm, showed a pattern of significance extending to the entire chromosome 10 with highest significance on distal end of the short arm and decreasing significance on the distal end of the long arm. The difference in the significance on the long arm between population IV and VI may due to the difference in population sizes, since the larger population VI is more powerful statistically in detecting skewing in segregation ratios.

Only one chromosomal region was identified in all three populations as being significantly distorted and associated with one selection model. However, chromosomal regions identified in

population VI may not have been identified in population III and/or IV due to the effect of their small sample sizes, since missing data and non-random selection would have a large effect on segregation ratios of small populations. Selections which are affected by environment and/or based on more than one genetic factors may also account for some of the complex segregation distortion across populations. In some cases of mixed patterns of significance for segregation distortion, combinations of selections in different developmental stages could be involved. Marker loci ordered with lower confidence in the linkage map may also cause mixed patterns of significance, since patterns of significance is based on the order of the marker loci derived from linkage analysis.

Segregation distortions observed in F_2 populations could result from selection occurring in pre-zygotic stages, including sporogenesis, gametic function and fertilization, and/or the post-zygotic stages including embryo and seed development, seed germination and plants growth (Zamir and Tadmor 1986). Since no reproductive barriers were observed in the fertility of the F_1^{Le} and F_1^{Lp4} plants which produced the F_2 populations (Liu et al. in press), it is not likely that segregation distortion in the F_2 populations is caused by selection in sporogenesis. In vivo pollen tube growth following self-pollinations of the F_1^{Le} and F_1^{Lp4} plants was normal, and readily sets seed (Liu et al. in press). Therefore, segregation distortions observed in our F_2 populations may result from selections occurring during fertilization or thereafter. Since some gametic selection models and zygotic selection models share similar patterns of significance, the models of some cases of selection can not be distinguished based on segregation ratios alone. Methods enable scoring the genotypes of gametes and zygotes in different developmental stages would provide more direct evidence for identifying the mechanism of segregation distortion. One example is screening the gene product of a reporter gene which is linked to a locus under selection and expresses in the developmental stages of interest. However, highly delicate microscopic techniques are needed for scoring the genotypes of gametes and zygotes in many developmental stages.

Segregation distortion could have a major impact on the results obtained from contemporary genome analyses, since the existence of distortion would violate the assumption in most of their protocols of Mendelian segregation for all loci in the genome. For example, the determination of LOD thresholds for QTL (quantitative trait loci) mapping (Churchill and Doerge 1994) was affected by the distortion detected on chromosome 10 (Mutschler, submitted), resulting in a higher probability of locating a false QTL by chance in that chromosome than in rest of the genome. Therefore, a better understanding of segregation distortion will be useful in adjusting methods of genome analyses.

Segregation distortion and hybrid breakdown, which involves various degrees of male sterility or non-fecundity in progeny of interspecific crosses, have both been observed in the progeny of interspecific hybrids from the genera *Gossypium*, *Lycopersicon*, and *Cucurbita* (Stephens 1950; Rick 1963; Rick 1969; Weeden and Robinson 1986). Coordinated studies of segregation distortion and hybrid breakdown could determine whether there is a relationship between this two phenomena. Such information could provide better understanding of the mechanism(s) of segregation distortion and interspecific crossing barriers and means to overcome these obstacles in crop improvement programs.

The authors thank Edward D. Cobb for the assistance in sample collection and greenhouse management and Dr. Nai-Syin Wang for helpful comments. This research was supported by Hatch Project 149484 and DOE Competitive Grant No. DE-FG02-91ER20037.

References

Agresti A (1990) Inference for two-way contingency tables. In: A Agresti (ed) Categorical Data Analysis. John Wiley & Sons, Inc., New York, NY, pp 36-78

Bernatzky R, Tanksley SD (1986) Majority of random cDNA clones correspond to single loci in the tomato genome. Mol Gen Genet 203:8-14

Bonierbale MW, Plaisted RL, Tanksley SD (1988) RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics 120:1095-1103

Chetelat RT, DeVerna JW (1991) Expression of unilateral incompatibility in pollen of *Lycopersicon pennellii* is determined by major loci on chromosomes 1,6 and 10. Theor Appl Genet 82:704-712

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963-971

de Vicente MC (1991) Genetic Studies in a Wide Cross of Tomato Using Restriction Fragment Length Polymorphisms (RFLP): Heterosis, Transgression, Sex-Specific Recombination. PhD Diss, Cornell University, Ithaca, NY

de Vicente MC, Tanksley SD (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

Doyle JJ, Dickson EE (1987) Preservation of plant samples for DNA restriction endonuclease analysis. Taxon 36:715-722

Gadish I, Zamir D (1987) Differential zygotic abortion in an interspecific *Lycopersicon* cross. Genome 29:156-159

Kinzer SM (1989) RFLP Mapping of Ripening Specific or Related cDNA Clones from *Lycopersicon esculentum* (tomato). MS Diss, Cornell University, Ithaca, NY

Kinzer S, Schwager S, Mutschler MA (1990) Mapping of ripening-related or -specific cDNA clones of tomato (*Lycopersicon esculentum*). Theor Appl Genet 79:489-496

Lincoln SE, Daly MJ, Lander ES (1993) Constructing Genetic Linkage Maps with MAPMAKER/EXP version 3.0: A Tutorial and Reference Manual. Research Technical Report 3rd edition

Lincoln SE, Lander ES (1992) Systematic detection of errors in genetic linkage data. Genomics 14:604-610

Mutschler MA (1990) Transfer of *Lycopersicon pennellii* into tomato (L. esculentum) does not create cytoplasmic male sterility. Rpt Tom Genet Coop 40:25-26

Rick CM (1963) Differential zygotic lethality in a tomato species hybrid. Genetics 48:1497-1507

Rick CM (1969) Controlled introgression of chromosomes of *Solanum pennellii* into *Lycopersicon esculentum*: segregation and recombination. Genetics 62:753-768

Rick CM (1972) Further studies on segregation and recombination in backcross derivatives of a tomato species hybrid. Biologisches Zentralblatt 91:209-220

Sato Y, Ishikawa R, Morishima H (1990) Nonrandom association of genes and characters found in *indica* x *japonica* hybrids of rice. Heredity 65:75-79

Stephenson JC (1991) The effect of gamma irradiation on protoplast fusion between atrazine-resistant *Brassica campestris* and ogura male sterile *B. oleraceae.* MS Diss, Cornell University, Ithaca, NY

Tanksley SD, Ganal MW, Prince JP, de Vincente MC, Bonierbale MW, Broun P, Fulton FM, Giovannoni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Roder MS, Wing RA, Wu W, Young ND (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics 132:1141-1160

Wagner H, Weber WE, Wricke G (1992) Estimating linkage relationship of isozyme markers and morphological markers in sugar beet (*Beta vulgaris* L.) including families with distorted segregations. Plant Breeding 108:89-96

Weeden NF, Robinson RW (1986) Allozyme segregation ratios in the interspecific cross *Cucurbita maxima* and *C. ecuadorensis* suggest that hybrid breakdown is not caused by minor alterations in chromosome structure. Genetics 114:593-609

Young ND, Tanksley SD (1989) Restriction fragment length polymorphism maps and the concept of graphical genotypes. Theor Appl Genet 77:95-101

Zamir D, Tadmor Y (1986) Unequal segregation of nuclear genes in plants. Bot Gaz 147:355-358

Table 1. Comparison of the G^2 tests for monogenic segregation ratios of the molecular markers that exhibited segregation distortion in populations III and IV.

		Population III (F_2^{Le})					Population IV (F_2^{Lp4})				
		G^2 test[c]					G^2 tests[c]				
marker[a]	chr[b]	1:2:1	1:1	e^2:ep:p^2	p(E)[d]	p(EP)[d]	1:2:1	1:1	e^2:ep:p^2	p(E)[d]	p(EP)[d]
TG224	1	*	-	*	0.43	0.65	no segregation				
TG19	1	*	-	*	0.47	0.69	-	-	-	0.48	0.47
TG430B	1	-	-	-	0.54	0.53	-	-	*	0.51	0.34
CD33A	2	-	-	-	0.43	0.56	-	-	*	0.50	0.65
TG469	2	-	-	-	0.46	0.57	*	-	**	0.45	0.68
TG48	2	-	*	-	0.38	0.51	*	-	*	0.43	0.64
TG141A	2	-	-	-	0.53	0.56	-	-	*	0.48	0.64
TG130	3	-	*	-	0.39	0.45	*	-	*	0.43	0.64
TG235	3	**	***	-	0.34	0.44	****	***	*	0.35	0.58
CT90A	3	*	*	-	0.35	0.45	-	-	-	0.40	0.56
TG134	3	*	*	-	0.38	0.48	-	-	-	0.45	0.56
TG94	3	-	-	*	0.51	0.64	-	-	-	0.50	0.52
TG272B	4	*	*	-	0.37	0.56	-	-	-	0.46	0.52
CD64	5	-	-	*	0.45	0.32	-	-	-	0.48	0.45
TG379	5	-	-	*	0.48	0.33	-	-	-	0.45	0.62
TG318	5	-	-	*	0.47	0.34	-	-	-	0.47	0.57
TG470	6	-	-	-	0.50	0.49	*	-	*	0.41	0.63
CT90B	6	-	-	*	0.47	0.67	-	-	-	0.46	0.56
TG115	6	-	-	-	0.49	0.50	-	-	*	0.50	0.36
TG166A	7	-	*	-	0.38	0.52	-	-	-	0.48	0.61
TG61	7	-	*	-	0.37	0.47	-	-	-	0.45	0.53
TG272A	7	-	-	*	0.49	0.69	-	-	-	0.48	0.60
TG18	9	**	**	-	0.36	0.58	-	-	-	0.47	0.46
TG223A	9	-	*	-	0.39	0.48	-	-	-	0.48	0.51
TG35	9	*	*	-	0.38	0.51	-	-	-	0.46	0.55
TG230	10	****	****	-	0.26	0.46	****	****	-	0.30	0.45
CT234	10	****	****	*	0.33	0.57	****	****	-	0.30	0.49
TG43	10	*	**	-	0.36	0.52	**	***	-	0.36	0.40
CD72	10	****	****	-	0.27	0.40	-	-	-	0.42	0.57
TG63	10	***	***	-	0.35	0.33	-	-	-	0.49	0.59
TG194	11	*.	**	-	0.37	0.42	-	-	-	0.43	0.43
CT211A	12	**	-	***	0.48	0.71	-	-	-	0.45	0.55
Pgi-1	12	-	*	-	0.39	0.47	-	-	-	0.47	0.55
TG111	12	*	*	-	0.36	0.38	-	-	-	0.48	0.69
TG28A	12	-	*	-	0.40	0.50	-	-	*	0.49	0.65

[a] The molecular markers showed significant segregation distortion in populations III or IV. Vertical line indicates adjacent marker loci.

[b] chr: Chromosome number

[c] Significance of the 1:2:1 G^2 test (1:2:1), 1:1 G^2 test (1:1), and e^2:2ep:p^2 G^2 test (e^2:2ep:p^2) for the markers in each population are presented as: $p \geq 0.05$ ('-' = not significant), $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.005$ (***), and $p < 0.001$ (****).

[d] p(E) and p(EP) are the frequencies of the E allele and the EP genotype, respectively, in the populations.

Table 2. Associations of gametic selection at an *A* locus with the heterozygote frequencies and significance of the G^2 tests for the segregation ratios of a *B* locus (a marker locus linked to the *A* locus).

models[a]	*p*, *q*[b]	magnitude of affecting factors[c]	p(EP)[d]	G^2 tests[e] 1:2:1	1:1	e^2:2ep:p^2
A	$pq>0$	+	< 50%	-	sig	-
A	$pq>0$	+++	< 50%	sig	sig	-
B	$pq<0$, $\lvert p\rvert \approx \lvert q\rvert$	+	> 50%	-	-	sig
B	$pq<0$, $\lvert p\rvert \approx \lvert q\rvert$	+++	> 50%	sig	-	sig
B'	$pq<0$, $\frac{1}{2}<\lvert\frac{p}{q}\rvert$ and $\lvert\frac{q}{p}\rvert<1$	+	> 50%	-	-	sig
B'	$pq<0$, $\frac{1}{2}<\lvert\frac{p}{q}\rvert$ and $\lvert\frac{q}{p}\rvert<1$	++	> 50%	sig	-	sig
B'	$pq<0$, $\frac{1}{2}<\lvert\frac{p}{q}\rvert$ and $\lvert\frac{q}{p}\rvert<1$	+++	> 50%	sig	sig	sig
B"	$pq<0$, $\lvert\frac{p}{q}\rvert$ or $\lvert\frac{q}{p}\rvert<\frac{1}{2}$	+	> 50%	-	sig	-
B"	$pq<0$, $\lvert\frac{p}{q}\rvert$ or $\lvert\frac{q}{p}\rvert<\frac{1}{2}$	++	> 50%	sig	sig	-
B"	$pq<0$, $\lvert\frac{p}{q}\rvert$ or $\lvert\frac{q}{p}\rvert<\frac{1}{2}$	+++	> 50%	sig	sig	sig
C	only $\lvert p\rvert \approx 0$ or only $\lvert q\rvert \approx 0$	+	≈50%	-	sig	-
C	only $\lvert p\rvert \approx 0$ or only $\lvert q\rvert \approx 0$	++	≈50%	sig	sig	-
C	only $\lvert p\rvert \approx 0$ or only $\lvert q\rvert \approx 0$	+++	≈50%	sig	sig	sig

[a] A: selection against the same gamete genotype in both sexes
B, B', B": selection against different gamete genotypes in the two sexes
C: selection against one gamete genotype in only one sex

[b] p : the proportion of the "A" genotype being eliminated in the female gametes
q : the proportion of the "A" genotype being eliminated in the male gametes
$-1 < p, q < 1$ ($-1 < p, q < 0$ indicating "a" gamete being eliminated)

[c] +: weak selection force and/or small population size and/or *B* far from *A* locus
++: moderate selection force and/or medium population size and/or *B* moderately distant from *A* locus
+++: strong selection force and/or large population size and/or *B* near *A* locus

[c] p(EP): frequency of the EP genotype

[e] sig: p-value < 0.05, -: non-significant

Table 3. Associations of zygotic selection at an *A* locus with the heterozygote frequencies and significance of the G^2 tests for the segregation ratios of a *B* locus (a marker locus linked to the *A* locus).

models[a]	*x*, *y*, *z*[b]	magnitude of affecting factors[c]	p(EP)[d]	G^2 goodness-of-fit tests[e] 1:2:1	1:1	e^2:2ep:p^2
D(dominant)	$x \approx y > 0$, $z \approx 0$	+	< 50%	-	sig	-
D(dominant)	$x \approx y > 0$, $z \approx 0$	+ +	< 50%	sig	sig	-
D(dominant)	$x \approx y > 0$, $z \approx 0$	+ + +	< 50%	sig	sig	sig
E(recessive)	$x \approx y \approx 0$, $z > 0$	+	> 50%	-	sig	-
E(recessive)	$x \approx y \approx 0$, $z > 0$	+ +	> 50%	sig	sig	-
E(recessive)	$x \approx y \approx 0$, $z > 0$	+ + +	> 50%	sig	sig	sig
F(overdominant)	$x \approx z \approx 0$, $y > 0$	+	< 50%	-	-	sig
F(overdominant)	$x \approx z \approx 0$, $y > 0$	+ + +	< 50%	sig	-	sig
G(underdominant)	$x \approx z > 0$, $y \approx 0$	+	> 50%	-	-	sig
G(underdominant)	$x \approx z > 0$, $y \approx 0$	+ + +	> 50%	sig	-	sig
H(codominant)	$x > y > 0$, $z \approx 0$	+	$\lessgtr$ 50%	-	sig	-
H(codominant)	$x > y > 0$, $z \approx 0$	+ +	$\lessgtr$ 50%	sig	sig	-
H(codominant)	$x > y > 0$, $z \approx 0$	+ + +	$\lessgtr$ 50%	sig	sig	sig

[a] D(dominant): selection against "AA" and "Aa" genotypes
E(recessive): selection against "aa" genotype
F(overdominant): selection against "Aa" genotype
G(underdominant): selection against "AA" and "aa" genotypes
H(codominant): selection against "AA" and "Aa" genotypes

[b] x : the proportion of the "AA" genotype being eliminated
y : the proportion of the "Aa" genotype being eliminated
z : the proportion of the "aa" genotype being eliminated
$0 < x, y, z < 1$

[c] +: weak selection force and/or small population size and/or *B* far from *A* locus
++: moderate selection force and/or medium population size and/or *B* moderately distant from *A* locus
+++: strong selection force and/or large population size and/or *B* near *A* locus

[d] p(EP): frequency of the EP genotype

[e] sig: p-value < 0.05, -: non-significant

Table 4. Comparison of the significance of the G^2 tests for monogenic segregation ratios of the RFLP markers located on chromosome 10.

	Kinzer et al 1990[b]					de Vicente 1991[b]					Chetelat and DeVerna 1991[b]				
Marker[a]	1:2:1	1:1	e^2:2ep:p^2	p(E)	p(EP)	1:2:1	1:1	e^2:2ep:p^2	p(E)	p(EP)	1:2:1	1:1	e^2:2ep:p^2	p(E)	p(EP)
TG230	n	n	n	n	n	****	****	****	0.27	0.46	n	n	n	n	n
TG122	****	****	****	0.30	0.57	n	n	n	n	n	****	****	*	0.26	0.49
CD56	****	****	***	0.33	0.60	n	n	n	n	n	n	n	n	n	n
TG303	n	n	n	n	n	****	****	***	0.29	0.46	n	n	n	n	n
CD45	***	*	*	0.40	0.61	n	n	n	n	n	n	n	n	n	n
TG560	n	n	n	n	n	****	****	*	0.30	0.46	n	n	n	n	n
TG43	-	*	-	0.42	0.56	n	n	n	n	n	n	n	n	n	n
TG103	**	*	*	0.41	0.63	n	n	n	n	n	***	****	-	0.35	0.46
TG280	n	n	n	n	n	****	****	-	0.31	0.44	n	n	n	n	n
CD34B	-	-	-	0.42	0.51	n	n	n	n	n	n	n	n	n	n
TG420	n	n	n	n	n	****	****	-	0.32	0.47	n	n	n	n	n
CD72	-	-	-	0.43	0.56	n	n	n	n	n	n	n	n	n	n
TG229	n	n	n	n	n	****	****	****	0.36	0.55	**	***	-	0.36	0.48
TG63	n	n	n	n	n	****	****	-	0.42	0.50	n	n	n	n	n
CD5	*	*	-	0.41	0.41	n	n	n	n	n	n	n	n	n	n
CD32B	-	-	-	0.45	0.49	***	****	-	0.44	0.50	n	n	n	n	n

[a] Markers are listed in the orders of the linkage groups of the publications. Relative order of markers from different publications were determined base on the tomato linkage map published by Tanksley et al (1992). Bold markers were also analyzed in populations III, IV, and VI.

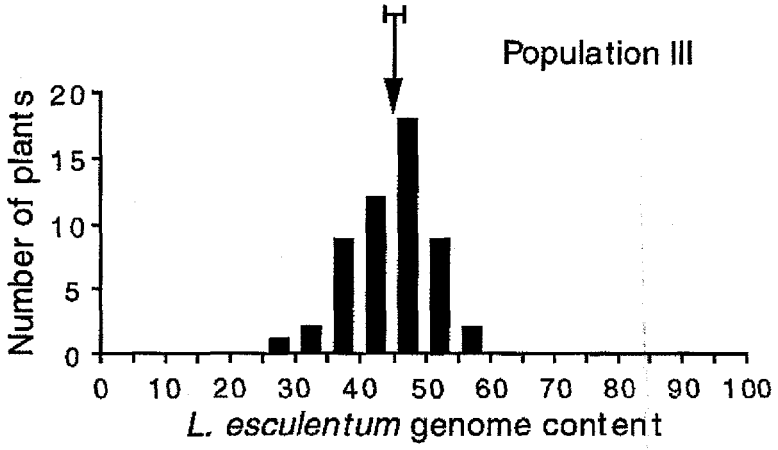
[b] Significance of the 1:2:1 G^2 test (1:2:1), 1:1 G^2 test (1:1), and e^2:2ep:p^2 G^2 test (e^2:2ep:p^2) for the markers in each population are presented as: $p \geq 0.05$ ('-' = not significant), $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.005$ (***), and $p < 0.001$ (****). n: data not available. p(E) and p(EP): frequencies of the E allele and the EP genotype, respectively, in the populations

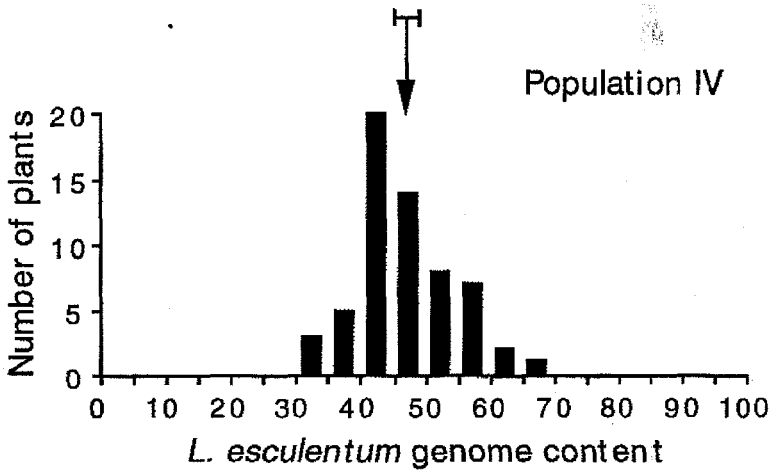
Figure 1. Frequency distributions of the *L. esculentum* genome content of populations III, IV, and VI. The arrows point to the population means, and the horizontal lines attached to the arrows indicate the 95% confidence intervals of the means.

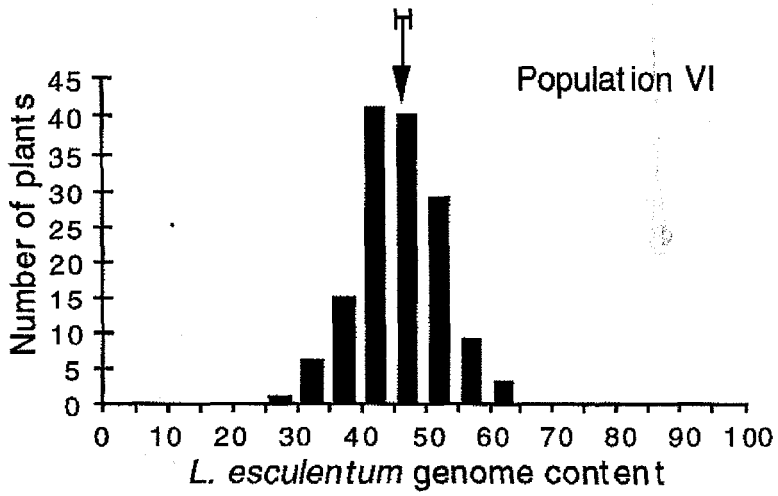
Figure 2. Frequency distributions of genome percentage of the EE (%EE), EP (%EP) and PP (%PP) genotypes of populations III, IV, and VI. The arrows point to the population means, and the horizontal lines attached to the arrows indicate the 95% confidence intervals of the means.

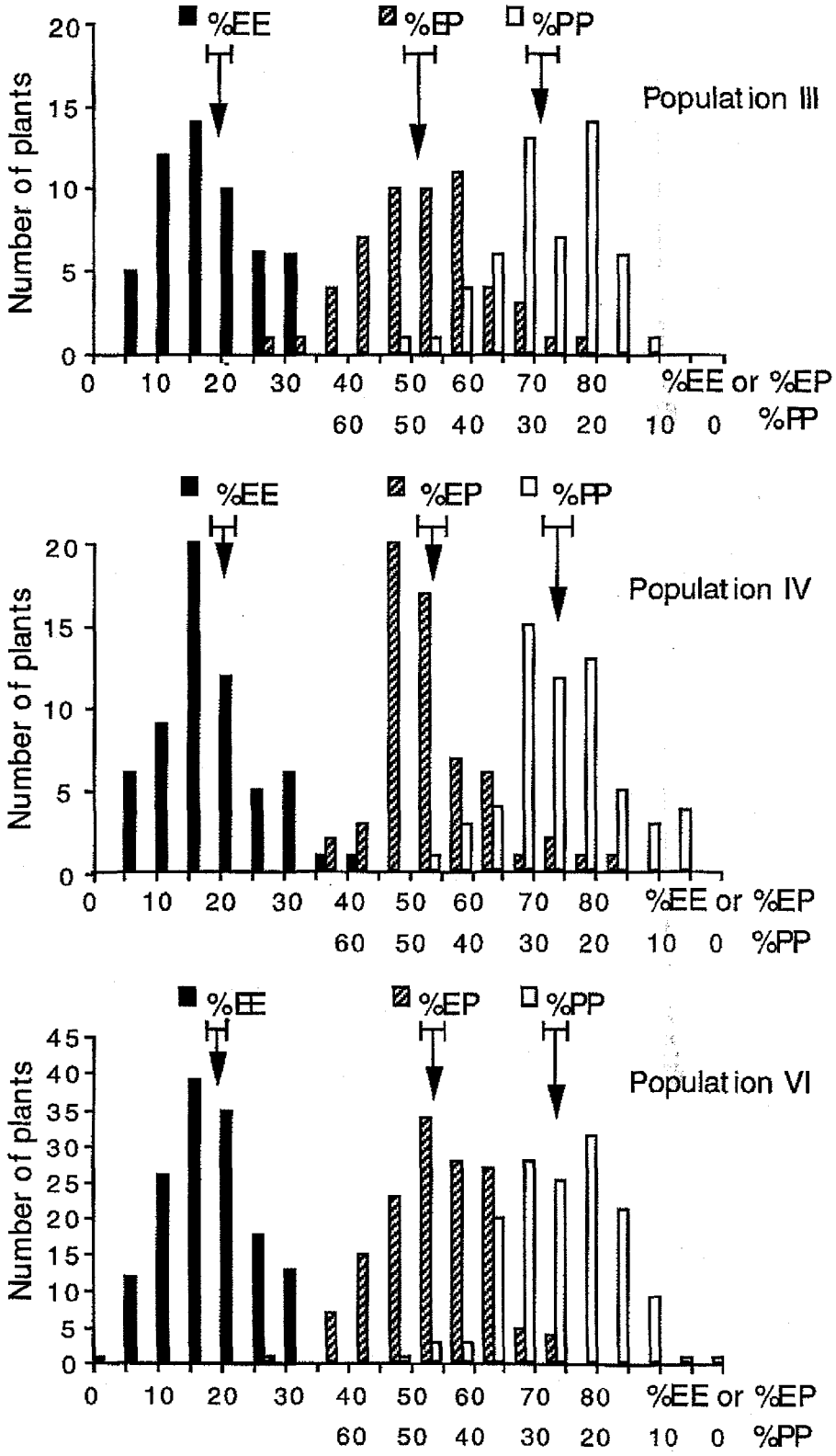
Figure 3. Triangle diagrams showing the genome percentage of EE (%EE), EP (%EP) and PP (%PP) genotypes of plants in population III, IV, VI, and a simulated population that has 150 plants generated from normal distributions with standard errors of 8.0 (the mean standard error of the percentages of EE, EP, and PP genome of population VI) and means of 25%, 50%, and 25%, respectively.

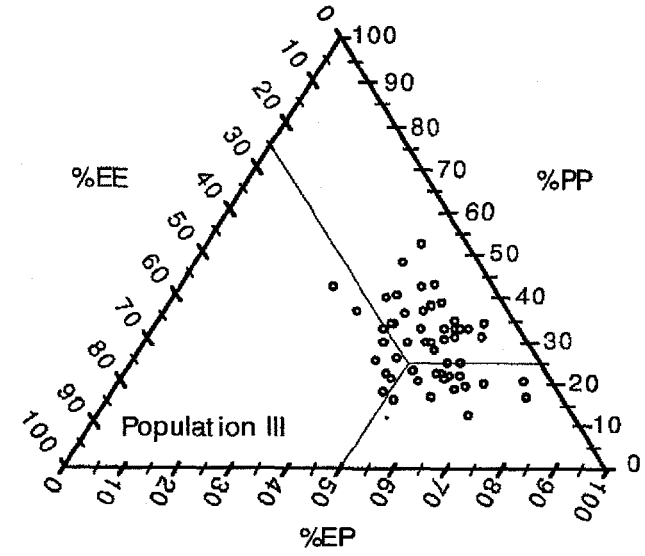
Figure 4. Genetic linkage maps of chromosomes showing segregation distortions identified by results of the G^2 goodness-of-fit tests for monogenic segregation ratios in population VI. Loci by tick marks were ordered with LOD > 3. Underlined loci were placed into the linkage group with LOD < 3.0. Loci following commas cosegregate. Significance of the 1:2:1 G^2 test (1:2:1), 1:1 G^2 test (1:1), and e^2:2ep:p^2 G^2 test (e^2:2ep:p^2) for the markers in each population are presented as: $p \geq 0.05$ ('-' = not significant), $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.005$ (***), and $p < 0.001$ (****). The p(E) and p(EP) show the frequencies of the E allele and the EP genotype, respectively, in population VI. Figure does not show the isolated significant marker loci in population VI, which including TG22 on chromosome 4 (significance of 1:2:1, 1:1, e^2:2ep:p^2 G^2 test: *,*,-), CD61 on chromosme 7 (significance of 1:2:1, 1:1, e^2:2ep:p^2 G^2 test: -,-,*), and TG28B and CD2 on chromosome 12 (significance of 1:2:1, 1:1, e^2:2ep:p^2 G^2 test: -,-,*, and -,*,-, respectively).

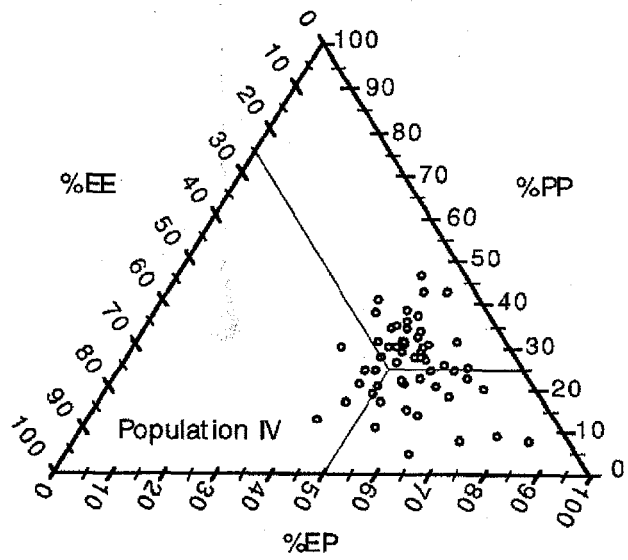

Population III
Number of plants
0
5
10
15
20
0 10 20 30 40 50 60 70 80 90 100
L. esculentum genome content

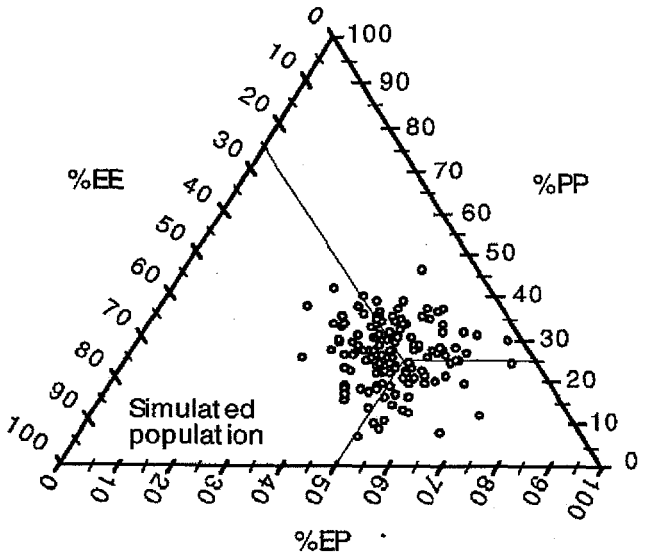

Population IV
Number of plants
0
5
10
15
20
0 10 20 30 40 50 60 70 80 90 100
L. esculentum genome content

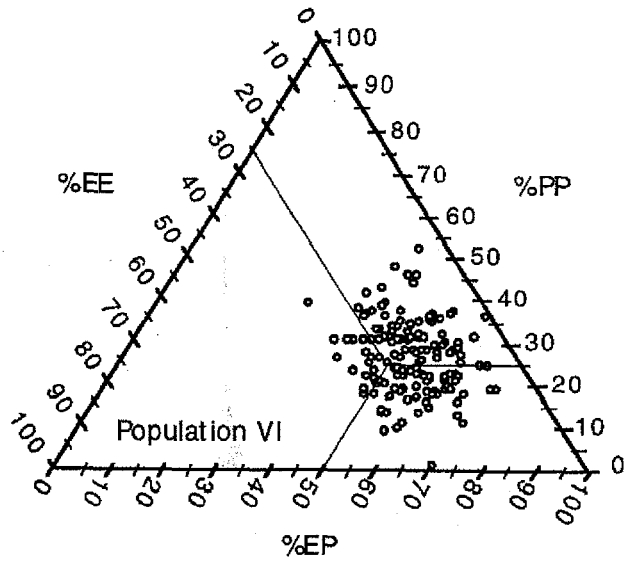

Population VI
Number of plants
0
5
10
15
20
25
30
35
40
45
0 10 20 30 40 50 60 70 80 90 100
L. esculentum genome content


%EE
%EP
%PP
Population III
Number of plants
0
5
10
15
20
0 10 20 30 40 50 60 70 80 %EE or %EP
60 50 40 30 20 10 0 %PP
%EE
%EP
%PP
Population IV
Number of plants
0
5
10
15
20
0 10 20 30 40 50 60 70 80 %EE or %EP
60 50 40 30 20 10 0 %PP
%EE
%EP
%PP
Population VI
Number of plants
0
5
10
15
20
25
30
35
40
45
0 10 20 30 40 50 60 70 80 %EE or %EP
60 50 40 30 20 10 0 %PP


%EE
%PP
%EP
Population III


%EE
%PP
%EP
Population IV


%EE
%PP
%EP
Simulated
population


%EE
%PP
%EP
Population VI