Copyright © 1991 by the Genetics Society of America

Mendelian Factors Underlying Quantitative Traits in Tomato: Comparison Across Species, Generations, and Environments

Andrew H. Paterson,*,[1] Susan Damon,[†] John D. Hewitt,[†,‡] Daniel Zamir,[§] Haim D. Rabinowitch,[§] Stephen E. Lincoln, Eric S. Lander,**,[††] and Steven D. Tanksley*,[2]**

**Department of Plant Breeding and Biometry, Cornell University, Ithaca, New York 14853, [†]Department of Vegetable Crops, University of California, Davis, California 95616, [‡]Northrup-King Research, Gilroy, California 95021, [§]Department of Field and Vegetable Crops, Hebrew University of Jerusalem, Rehovot 76-100, Israel, **Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, and [††]Harvard University, Cambridge, Massachusetts 02138*

Manuscript received May 21, 1990
Accepted for publication September 14, 1990

ABSTRACT

As part of ongoing studies regarding the genetic basis of quantitative variation in phenotype, we have determined the chromosomal locations of quantitative trait loci (QTLs) affecting fruit size, soluble solids concentration, and pH, in a cross between the domestic tomato (*Lycopersicon esculentum* Mill.) and a closely-related wild species, *L. cheesmanii*. Using a RFLP map of the tomato genome, we compared the inheritance patterns of polymorphisms in 350 F_2 individuals with phenotypes scored in three different ways: (1) from the F_2 progeny themselves, grown near Davis, California; (2) from F_3 families obtained by selfing each F_2 individual, grown near Gilroy, California (F_3-CA); and (3) from equivalent F_3 families grown near Rehovot, Israel (F_3-IS). Maximum likelihood methods were used to estimate the approximate chromosomal locations, phenotypic effects (both additive effects and dominance deviations), and gene action of QTLs underlying phenotypic variation in each of these three environments. A total of 29 putative QTLs were detected in the three environments. These QTLs were distributed over 11 of the 12 chromosomes, accounted for 4.7–42.0% of the phenotypic variance in a trait, and showed different types of gene action. Among these 29 QTLs, 4 were detected in all three environments, 10 in two environments, and 15 only in a single environment. The two California environments were most similar, sharing 11/25 (44%) QTLs, while the Israel environment was quite different, sharing 7/20 (35%) and 5/26 (19%) QTLs with the respective California environments. One major goal of QTL mapping is to predict, with maximum accuracy, which individuals will produce progeny showing particular phenotypes. Traditionally, the phenotype of an individual alone has been used to predict the phenotype of its progeny. Our results suggested that, for a trait with low heritability (soluble solids), the phenotype of F_3 progeny could be predicted more accurately from the genotype of the F_2 parent at QTLs than from the phenotype of the F_2 individual. For a trait with intermediate heritability (fruit pH), QTL genotype and observed phenotype were about equally effective at predicting progeny phenotype. For a trait with high heritability (mass per fruit), knowing the QTL genotype of an individual added little if any predictive value, to simply knowing the phenotype. The QTLs mapped in the *L. esculentum* × *L. cheesmanii* F_2 appear to be at similar locations to many of those mapped in a previous cross with a different wild tomato (*L. chmielewskii*). One possible explanation of this similarity is that genetic factors at some of the same loci may affect the traits in the two distantly-related wild species. Potentially major implications of such similarity across broad genetic distances are discussed, in regard to plant and animal breeding, germplasm introgression, and cloning of QTLs.

CONTINUOUS variation in phenotype, typical of most traits in nature and agriculture, represents the collective action of many "polygenes" lying at different "quantitative trait loci" (QTLs), together with environmental effects (JOHANSSEN 1909; NILSSON-EHLE 1909; EAST 1915). Techniques exist for estimating the "effective number" (*e.g.*, the minimum number: LANDE 1981) of QTLs influencing a trait (WRIGHT 1968; LANDE 1981), and for estimating the average gene action (MATHER and JINKS 1971) of QTLs influencing a trait. However, these techniques are useful in describing average properties of a group of QTLs, rather than defining the location and specific phenotypic effects of individual QTLs. Early attempts to describe properties of individual QTLs, by studying linkage to easily-scored morphological mutations (first by SAX 1923, in greatest detail by THODAY 1961), were impaired by a lack of suitable markers in many genetic stocks, and by confounding effects of the markers themselves on the traits under study.

High density genetic linkage maps of DNA markers

[1] Current address: E. I. DuPont de Nemours, Wilmington, Delaware 19880-0402, and Department of Plant and Soil Sciences, University of Delaware, Newark, Delaware 19711.
[2] To whom reprint requests should be sent.

(mostly restriction fragment length polymorphisms, or RFLPs) permit one to determine the location of QTLs by linkage analysis (PATERSON *et al.* 1988, LANDER and BOTSTEIN 1989). Such maps, proposed only 10 years ago (BOTSTEIN *et al.* 1980), now exist for many plant and animal species (TANKSLEY *et al.* 1989). Genetic mapping with DNA markers utilizes preexisting DNA polymorphisms in a population, introduces no additional phenotypic variation into an experiment (unlike morphological mutations), and precludes the need to construct special "multiply marked stocks" for linkage analysis.

In a previous experiment, we mapped QTLs affecting mass per fruit, soluble solids concentration, and fruit pH, to ~20 cM regions of the genome, in a backcross of the wild Peruvian species *Lycopersicon chmielewskii* (CL) to the domestic tomato, *L. esculentum* (E) (PATERSON *et al.* 1988). This provided a detailed picture of the inheritance of these three quantitative traits, including the approximate locations and phenotypic effects of QTLs which collectively accounted for 44–58% of the phenotypic variation in these traits. In subsequent experiments, QTLs have been mapped to intervals of as little as 3 cM, narrowing the gap between linkage analysis and physical analysis of QTLs (PATERSON *et al.* 1990). This level of resolution appeared sufficient to separate some genetic factors causing undesirable wild traits from closely linked factors associated with elevated soluble solids concentration, a commercially valuable trait that we are attempting to transfer to the domestic tomato.

In the experiments described herein, we studied the genetic basis of quantitative variation in mass per fruit, soluble solids concentration, and fruit pH in a second wild species, *L. cheesmanii* (CM), from the Galapagos Islands. Although CL and CM are only distantly related (RICK 1979; MILLER and TANKSLEY 1990), both have much smaller fruit and much higher soluble solids concentration than E. Both CL (RICK 1974; OSBORN, ALEXANDER and FOBES 1987; TANKSLEY and HEWITT 1988) and CM (GARVEY and HEWITT 1984) are of considerable interest to breeders, for their high soluble solids concentration. These experiments were designed to investigate: (1) *Gene action of individual QTLs*. By using a F_2 population, all three possible gene dosages at a locus are represented. This permits estimation of additive effects and dominance deviations for individual QTLs. (2) *Predictive value of QTLs within a pedigree*, by phenotyping both F_2 plants and their self pollinated F_3 progeny, we investigated whether progeny phenotype could be predicted more accurately by parental phenotype or by parental QTL genotype. This involved comparing the observed phenotype in F_3 (family means) to that predicted from the F_2 phenotypes (by classical parent/offspring regression), and to that predicted from the F_2 QTL analysis (incorporating predictive value of QTLs). Further, this also explores what QTL mapping information might be extracted by testing self-pollinated progeny of a partially heterozygous mapping population, without the expense of additional RFLP genotyping in the lab (LANDER and BOTSTEIN 1989). (3) *Predictive value of QTLs across environments*, by comparing QTL maps of F_2 plants grown in Davis, California to F_3 families grown in California and sibling F_3 families grown in Israel, we investigated whether QTLs at similar chromosomal locations account for phenotypic variation in different environments. (4) *Predictive value of QTLs across evolutionary distances*, by comparing the QTL map of E × CM to a map of E × CL made previously (PATERSON *et al.* 1988), we investigated whether QTLs at similar chromosomal locations account for phenotypic variation in different species.

MATERIALS AND METHODS

Mating and field plot design: The populations studied were derived from crosses between *Lycopersicon esculentum* cultivar "UC204B" (hereafter E) and *L. cheesmanii* accession LA483 (hereafter CM). Both E and CM were homozygous at all marker loci examined. In the summer of 1987, 350 F_2 (self-pollinated progeny of a F_1 hybrid between E and CM) individuals were grown in the field in Davis, California, in a completely randomized design. In the summer of 1988, plots of up to 10 self-pollinated progeny from each F_2 individual were grown near Gilroy, California, and near Rehovot, Israel. Each of the F_3 experiments used a completely randomized design, however the 10 self-pollinated progeny of each F_2 individual were grown together in a plot. Because we studied both F_2 individuals and their self-pollinated F_3 progeny, each F_2 individual could in fact be assigned *three* different phenotypic values for a trait—one from the F_2 individual itself, a second from the average of the ten F_3 progeny grown in California, and a third from the average of the ten F_3 progeny grown in Israel.

Phenotyping: For each F_2 plant, mass per fruit, soluble solids concentration, and fruit pH were determined as described elsewhere (TANKSLEY and HEWITT 1988). In addition, the F_2 plants were scored visually for fruit color (red/orange/yellow) and determinacy (differentiation of the apical meristem into a flower cluster).

For each plant in each F_3 family, traits were measured in the same way as for the F_2 plants (with one exception) and then averaged across each F_3 family. The one exception is the following: In Rehovot, fruit *diameter* rather than mass per fruit was measured. Although these two traits are correlated, this difference slightly complicates the interpretation of results for the F_3 grown in Rehovot. We also note that measurements of F_3 fruit characteristics were performed at the Gilroy and Rehovot sites, respectively—any slight differences between laboratories could be confounded with environmental variation.

Genotyping: RFLP genotypes were determined as described in TANKSLEY and HEWITT (1988), except that probes were labeled by primer extension (FEINBURG and VOGELSTEIN 1983). The probes used were a subset of those previously mapped in a different tomato cross (BERNATZKY and TANKSLEY 1986; TANKSLEY *et al.* 1988), selected for coverage of the tomato genome at ~20 cM intervals. Each of the probes used showed polymorphism between E and CM

in genomic digests with at least one restriction enzyme (of up to 16 studied).

Data analysis: The following analyses were performed.

Trait means, correlations, and standard-unit parent/progeny regressions: Each of these were determined using SAS (SAS Institute 1988). Parent-progeny regressions were in standard units (FREY and HORNER 1957).

Narrow-sense heritabilities: These were calculated from the slope of the classical F_3-F_2 regression, adjusted for inbreeding (SMITH and KINMAN 1965).

Segregation ratios of individual markers: These were summarized using SAS (SAS Institute 1988), and chi-square values computed using a simple computer macro (A. H. PATERSON, unpublished results).

Linkage maps: These were constructed using MAPMAKER (LANDER *et al.* 1987). All comparisons of map lengths were based on recombination fraction. Standard errors for recombination fraction were determined according to ALLARD (1956).

Genome composition: This was estimated based on marker genotypes and the recombinational distances between markers, as previously described (PATERSON *et al.* 1988). Because the present study used a F_2 population, there were three possible genotypes for any interval (E, CM, or heterozygote), rather than just two (E, heterozygote) in the prior study. When consecutive markers along the chromosomes of an individual show the same genotype, the estimates assume that the region intervening between the markers is comprised entirely of the marker genotype. When consecutive markers along the chromosomes of an individual show a different genotype, the estimates assume that the interval is comprised of half each genotype. Such estimates disregard the possibility of double recombinants within an interval; on average, this has no net effect on estimating overall genome composition, although it would cause an underestimate of the total *number* of recombinant chromosome segments across the genome.

QTL likelihood maps: The chromosomal locations of putative QTLs were determined by the method of interval mapping (LANDER and BOTSTEIN 1989), and are represented (Figure 3) as QTL likelihood plots (PATERSON *et al.* 1988). We explain the method briefly here, since it requires extension of a previous method applied to backcross populations (LANDER and BOTSTEIN 1989).

For a model assuming m QTLs at hypothesized locations, phenotypes are assumed to be determined by an equation of the form:

$$\phi j = \sum_{i=1,\ldots,m}(aij\ xij + dij\ yij) + ej$$

where j is the phenotype of the jth individual, xij is 0, 1 or 2 according to whether the jth individual possesses 0, 1 or 2 copies of the CM allele at the ith QTL, yij is the 1 or 0 according to whether the jth individual is heterozygous or homozygous at the ith QTL and ej is a random normal variable with mean 0 and standard deviation σ. Thus, the coefficients aij, dij and ej represent additive effects, dominance effects and residual noise, respectively. Except when a QTL is located exactly at the position of a genetic marker, the QTL genotype is not known with certainty. Nevertheless, the probability distribution over the possible QTL genotypes can be determined based on the genotypes at the nearest flanking markers on either side of the m putative QTLs (with the probability based on the chance of various patterns of noncrossovers, single crossovers and double crossovers). Based on this distribution and a given value of the coefficients, one can calculate the probability (density) that the model would have given rise to the observed phenotypes of the population. By maximizing this probability

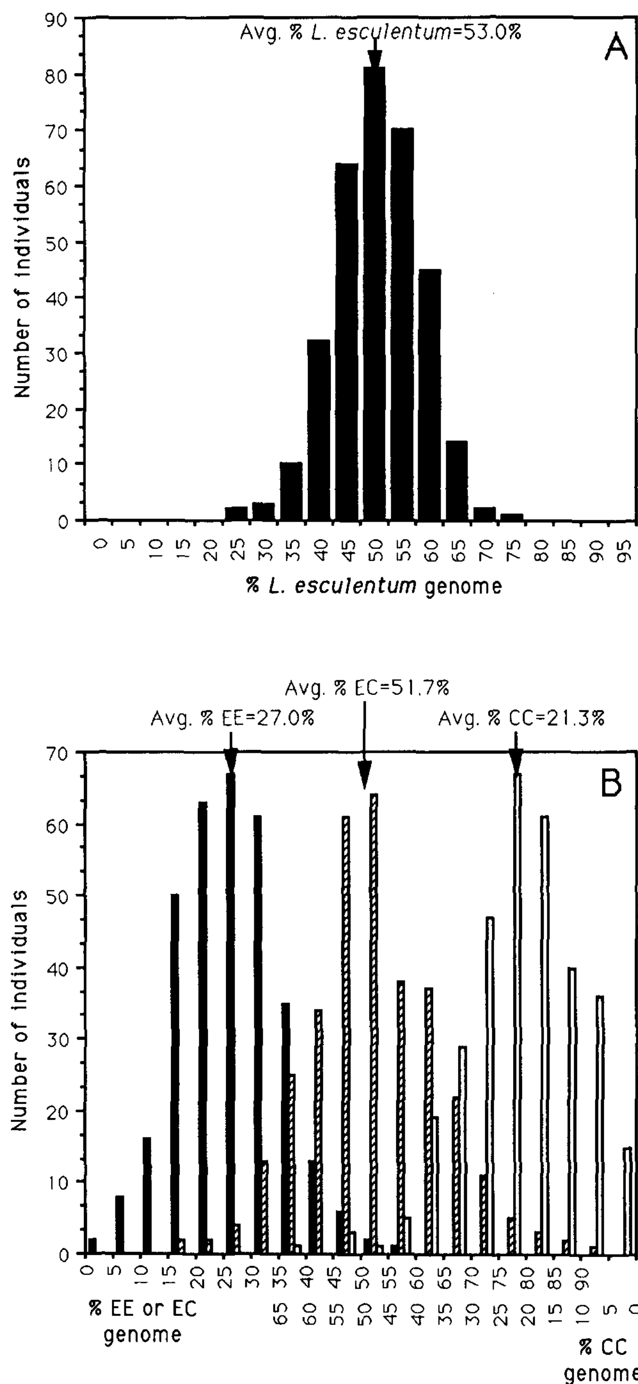

FIGURE 1.—(A) Frequency distribution for % domestic parent (E) genome, and (B) frequency distribution for percent heterozygosity (EC) and percent homozygosity for wild (CC) and domestic (EE) parent genomes, in the F_2 progeny of *L. esculentum* × *L. cheesmanii*, estimated as described in MATERIALS AND METHODS.

over the possible values for the coefficients, one can find the maximum likelihood estimates of aij, dij and σ. To compare models, one examines the ratio of the probability of the data arising under each model. This ratio is called the odds ratio and its $\log_{10}$ is called the LOD score (LANDER and BOTSTEIN 1989).

QTL likelihood maps are produced as follows. At each point of the genome, one computes the LOD score for the comparison of the hypothesis of a putative QTL at that position (fitting a single additive, dominance and noise term)

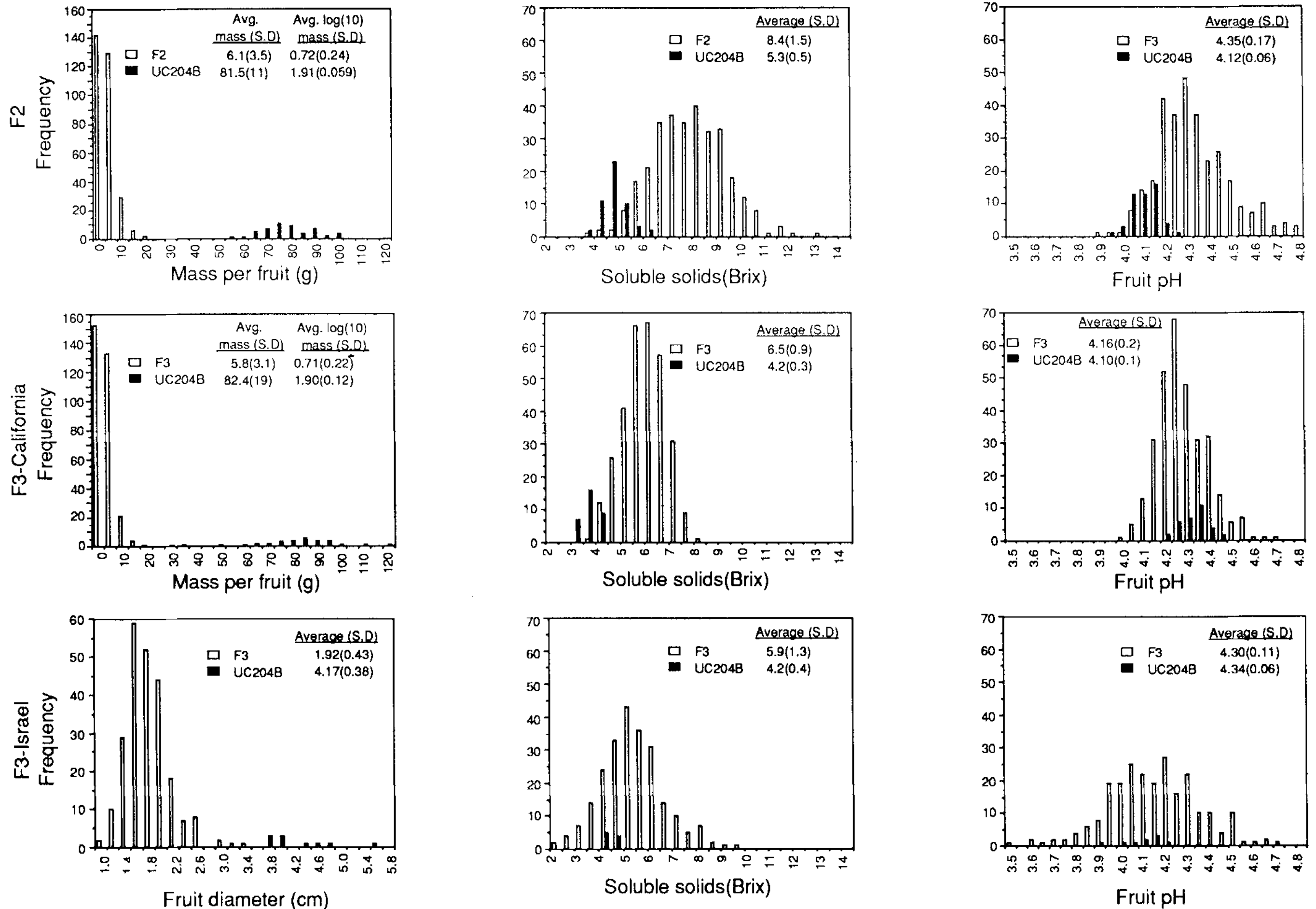

FIGURE 2.—Frequency distributions for mass per fruit, soluble solids concentration, and pH in the E parental strain (filled bars) and the E × CM F_2 progeny (open bars). The F_2 distributions for soluble solids concentration and pH are approximately normal, but mass per fruit is skewed; $\log_{10}$(mass per fruit) was used in quantitative analyses to improve normality.

with the hypothesis of no QTL (fitting only a noise term). Possible QTLs elsewhere in the genome are ignored (*e.g.*, included in the noise term) in this first-generation mapping. If the LOD score exceeds a predetermined threshold (see discussion below), a QTL is inferred to be present. The position of the QTL is estimated to be the interval over which the LOD score is within a certain threshold of the peak. Typically, such support intervals are defined with a threshold of 1 or 2 log-units, indicating the region over which the model's probability of giving rise to the data is at most 10- or 100-fold less than at the most likely position.

In QTL likelihood plots, the curves may appear to have multiple nearby peaks (Figure 3). This does not necessarily indicate the presence of multiple QTLs: if a QTL is actually present in one interval, the hypothesis of a QTL in an adjacent interval will still fit the data better than the hypothesis of no QTL at all, and the more likely position of a QTL in this adjacent interval will often be near the middle of the interval (since this position is furthest from any potentially conflicting data at the observed markers). Accordingly, multiple peaks correctly reflect the shape of the likelihood surface but need not indicate multiple QTLs. However, two peaks separated by considerable distance (say 50 cM) are likely to represent two distinct QTLs. To test the hypothesis of multiple QTLs, one can fix the position and effect of one QTL, then compute a QTL likelihood plot showing the LOD score for a two-QTL model; if the LOD score is significantly higher for two QTLs than for one QTL, the presence of two distinct QTLs is likely.

The gene action of individual QTLs was investigated by several steps. As described above, QTLs were first located by testing the hypothesis of a = 0 + d = 0 (the "unconstrained genetics" model) at 1 cM intervals across the entire genetic map. Then, we wished to determined whether the action of each QTL was largely additive, dominant, or recessive. This was done by evaluating the relative likelihood of models which constrained the QTL to particular "pure" types of gene action. Specifically, one tests a purely additive model by forcing the dominance term $d = 0$, a dominant model by forcing $d = a$, and a recessive model by forcing $d = -a$. A 1-LOD (10-fold) reduction in likelihood was considered to mean that the type of gene action was unlikely; the types of gene action indicated for QTLs are those which could not be rejected as unlikely (Figure 3, Tables 2–4). It is emphasized that failing to reject a type of gene action is not equivalent to demonstrating that the relevant gene exhibits that pure type of gene action. For example, for a particular QTL, rejection of the dominant and recessive models but not the additive model is still not sufficient evidence to assert that the effect of the QTL is "purely" additive. The choice of a 1-LOD likelihood difference for rejecting types of gene action is somewhat arbitrary, but is consistent with criterion used in human genetic models (OTT 1985). Such a likelihood ratio test is more appropriate than a simple analysis of variance test based on marker genotype(s) as the independent variable(s). The reason for this is

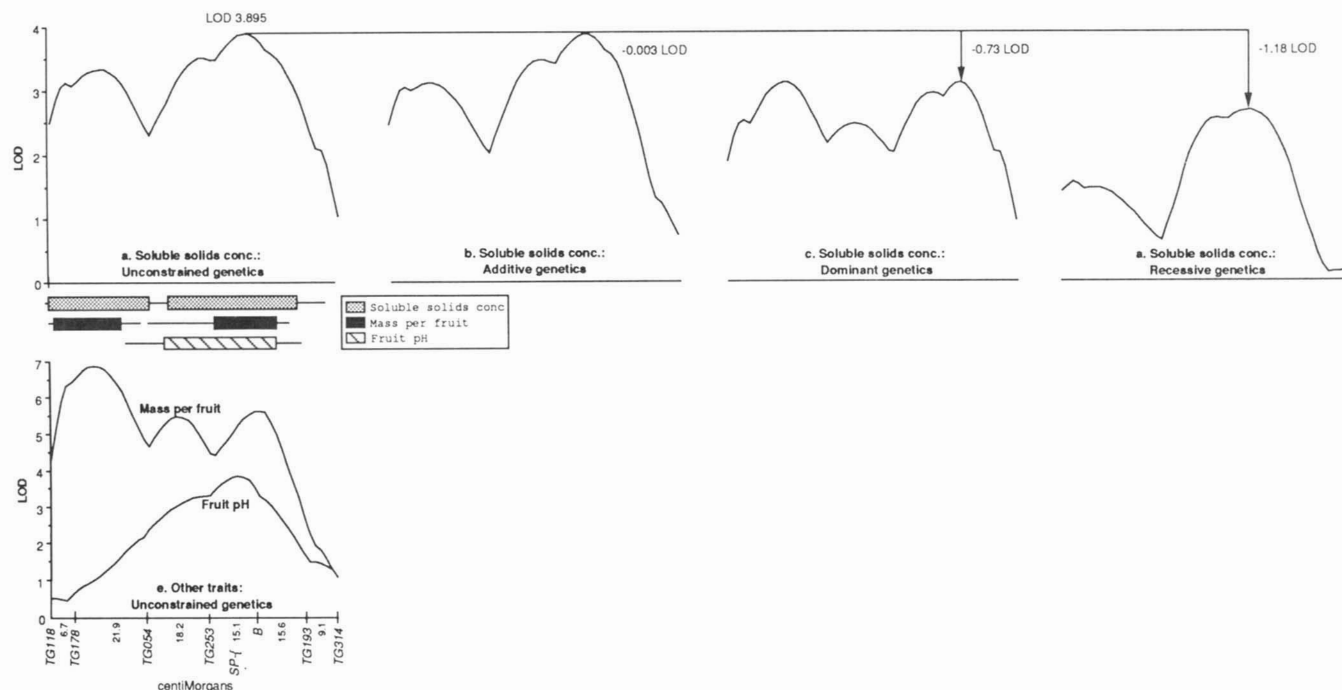

FIGURE 3.—QTL likelihood plots indicating LOD scores for soluble solids concentration (plots a–d), mass per fruit (e), and fruit pH (e), along chromosome 6. The RFLP linkage map used in the analysis is presented along the abscissa, in centiMorgans according to KOSAMBI (1944). The position of a QTL is shown as the interval over which the LOD score is within 1 or 2 log-units of that at the most likely position, indicating the region over which the model's probability of giving rise to the data is at most 10- or 100-fold less than at the most likely position. The 1-LOD (10-fold; bars) and 2-LOD (100-fold; whiskers) likelihood intervals for QTLs affecting each trait are presented between plots (a) and (e). After allowing for the QTLs near *B*, the residual LOD scores near *TG178* for mass per fruit and soluble solids were each sufficient (>2.0) to support a second QTL. The most likely type of gene action of a QTL was assessed by comparing the relative likelihoods (LOD scores) of additive, dominant, and recessive genetic models (as described in MATERIALS AND METHODS). For the indicated QTL (affecting soluble solids concentration), additive genetics (plot b) was most likely, dominant genetics (plot c) was less likely but could not be ruled out at the 1-LOD level, and recessive genetics (plot d) was unlikely by more than 1 LOD. Hence, the gene action of this locus is denoted AD in Tables 2–4, indicating that activity (A) is most likely but dominance (D) could not be falsified. All nomenclature for gene action (Tables 2–4) follows this system.

that whenever the QTL does not lie exactly at a marker locus (which is usually the case), recombination can occur between the marker and the QTL. Because of recombination, the marker genotype classes are actually mixtures of two distributions, a situation which violates an assumption of the analysis of variance. Our likelihood test accounts for this mixing due to recombination.

Finally, the appropriate significance threshold (LOD score) for declaring QTLs must be discussed. By using the mathematical theory of large deviations of normal processes, LANDER and BOTSTEIN (1989) have determined the appropriate threshold for significance for analyses involving a single degree of freedom at each locus. This applies to (1) QTL mapping in a backcross or recombinant inbred population, or (2) QTL mapping in an intercross when the alleles are constrained to have purely additive, purely recessive, or purely dominant phenotypic effects. By contrast, fitting both additive and dominance effects to QTLs in an intercross involves two degrees of freedom: the appropriate significance threshold depends on the corresponding mathematical theory of large deviations of generalized chi-squared processes, a much less developed subject. Because the solutions are not completely worked out (see LANDER and BOTSTEIN 1989 for approximations), we confirmed significance of all observed QTLs by evaluating the LOD for purely additive gene action. In this case, a threshold of 2.4 is appropriate (LANDER and BOTSTEIN 1989).

In addition to the likelihood interval and types of gene action (additive effects and dominance deviations), the percent of total phenotypic variation accounted for by each QTL has been determined. Values presented for each QTL (Tables 2–4) come from a model including only that QTL, using the genotype at the QTL (inferred from flanking RFLP markers) as the independent variable.

Further, the percent of phenotypic variation accounted for in each trait is presented (Table 5), from a model which includes the inferred genotypes at all QTLs with (individually) significant effects on the trait, as the independent variables.

Two-way interactions between loci: These were evaluated using the PROC GLM routine in SAS (SAS Institute 1988), using the genotypes at pairs of RFLP markers as the independent variables. The 4 d.f. for interlocus interaction were partitioned into additive * additive, additive * dominant, dominant * additive, and dominant * dominant epistasis, using four orthogonal contrasts.

RESULTS

Genome transmission

Segregation of markers: Deviation from the expected 1:2:1 genotype frequency was significant ($P \leq 0.05$) for 36 (51%) of the 71 markers scored, comprising 15 linked groups on 11 chromosomes, with only chromosome 5 segregating normally. Ten of

these 15 groups also showed significant deviation from the expected 1:1 allele frequency, seven favoring the E allele and three the CM allele. In two groups, deviation from 1:1 was indicated but fell slightly below the significance threshold. For the remaining three groups, the parental alleles occurred at similar frequency (nonsignificant deviation from 1:1), but the heterozygous genotype was overabundant (significant deviation from 1:2:1). Segregation distortion in the E × CM F_2 appeared to be somewhat less pronounced than in a backcross of E to the more distantly-related species *L. chmielewskii* (CL), which showed 21 regions of aberrant segregation (69% of markers), including parts of all 12 chromosomes.

The most extreme examples of segregation distortion in the E × CM F_2 occurred at CD35 on chromosome *2*, where only 5 of 298 plants scored were CM homozygotes, and CD71 on chromosome *3*, where only 13 of 310 plants scored were E homozygotes.

Recombination between markers: The E × CM F_2 linkage map based upon the 71 markers spans 1023 cM, with an average spacing of 17.3 cM between markers. The linear order of markers agrees with that found previously for several smaller populations of E × *L. pennellii* (P). One marker, CD17, fails to show linkage to the E × CM map, but has been mapped previously to one end of chromosome *11* (TANKSLEY *et al.* 1988). Lack of polymorphism between the closely related E and CM has left chromosomes *9* and *12* sparsely marked.

The E × CM and E × P F_2 maps show similar recombinational length, despite the fact that E and CM are much more closely related than E and P (RICK 1979; MILLER and TANKSLEY 199). While the two maps do not include all of the same markers, a subset of markers spanning 777 cM in the E × CM F_2 covers 760 cM of a E × P F_2 population, a nonsignificant difference. Both the E × CM and E × P F_2 populations show significantly greater recombination than a third population, a E × CL backcross, using the F_1 as male (PATERSON *et al.* 1988). Genetic distance between E and P appears to be at least as great as that between E and CL (MILLER and TANKSLEY 1990), and chromosome pairing in each of these wide crosses is fairly normal.

The higher recombination rates of the F_2 populations than the backcross may be due to greater reproductive success of recombinant gametes which are maternal. Recombinant gametes from either sex could be detected in the F_2 populations, while only paternal recombination was detectable in the backcross (since the F_1 was paternal). Higher recombination rates are found for maternal gametes in other tomato crosses (GADISH and ZAMIR 1987; C. VICENTE and S. D. TANKSLEY personal communication), as well as in maize (RHOADES 1941; ROBERTSON 1984), and in the homogametic sexes of Drosophila (MORGAN 1914), *Bombyx mori* (MAEDA 1939; TAZIMA 1964), *Mus musculus* (DAVISSON and RODERICK 1981), and *Homo sapiens* (RENWICK and SCHULZE 1965; DONIS-KELLER *et al.* 1987).

Genome composition of F_2 individuals: On average, the F_2 individuals were inferred to be homozygous for E alleles along 27.0% of the genome, heterozygous over 51.7% of the genome, and homozygous for CM across 21.3% of the genome (Figure 1A). In total, the genome composition of an average F_2 individual was inferred to be 53 (±7.7)% E (Figure 1B), very close to the expected 50%. The range of 26.4 – 78.1% E is close to the random expectation for a population of this size (e.g. ± about 3 standard deviations). We note that the extreme F_2 individuals carried about as much donor genome as an average backcross individual. As suggested elsewhere (TANKSLEY, MEDINA-FILHO and RICK 1981; PATERSON *et al.* 1988; YOUNG and TANKSLEY 1989), such variation in the proportion of recurrent parent genome can be used as a basis for genotypic selection among individuals. For example, one might accelerate the introgression (backcrossing) of a desired chromosomal segment into a new genetic background, by selecting progeny at each generation with a minimum of donor genotype. Alternately, one might accelerate inbreeding to homozygosity, by selecting progeny at each generation which are homozygous over a maximal proportion of the genome.

Phenotypic variation

Trait means: In the present studies, the E accession (UC204B) had mass per fruit of 81.5–82.4 g, while the CM accession (LA483) is characterized by mass per fruit of 3 g or less (S. D. TANKSLEY, unpublished results). This difference between E and CM represents about 20 SD, using $\log_{10}$-transformed data. The segregating generations had average mass per fruit of 6.1 (F_2) and 5.8 g (F_3-CA), both significantly less than E.

The soluble solids concentration of E ranged from 4.2–5.3 °Brix (Figure 2), about 10 SD less than the 14.4–16.5 °Brix reported for CM (GARVEY and HEWITT 1984). The segregating generations had average mass per fruit of 6.1 (F_2) and 5.8 g (F_3-CA), and soluble solids concentration of 8.4 (F_2), 6.5 (F_3-CA), and 5.9 (F_3-IS) °Brix, all significantly higher than E.

The pH of E ranged from pH 4.10–4.34, and the segregating generations were similar to E, with average pH of 4.35 (F_2), 4.16 (F_3-CA), and 4.30 (F_3-IS). Average pH of the F_2 was significantly higher than E, but neither of the F_3 averages were significantly different from E grown in the same environment. Reliable measurements of pH for CM were not available.

Fruit diameter (which, as noted above, was measured instead of mass per fruit in the Israel F_3 trial)

TABLE 1

Correlations among traits in F_2 and F_3 generations

Generation[a]	Solids × mass per fruit	Solids × pH	Mass per fruit × pH
F2	−0.59**	+0.27**	−0.21**
F3-CA	−0.06	+0.16**	−0.05
F3-IS	−0.25**	+0.16*	+0.08[b]
CLBC1	−0.42**	+0.33**	−0.08

*, ** denote significance at ≤0.05 and ≤0.01 levels, respectively.
[a] CA = Gilroy, California, IS = Rehovot, Israel, CLBC1 from PATERSON *et al.* (1988).
[b] Fruit diameter × pH.

averaged 1.92 cm for the F_3 plants, much smaller than the 4.17 cm average for E.

All of the traits studied showed continuous variation, typical of quantitative traits (Figure 2). Soluble solids concentration, fruit pH, and fruit diameter exhibited approximately normal distributions, while mass per fruit showed continuous variation but was skewed toward the small value of CM. The transformation $\log_{10}$(mass per fruit) improved normality, and was used in all analyses.

Trait correlations: Consistent with prior observations (GOLDENBERG and PAHLEN 1966; IBARBIA and LAMBETH 1969; PATERSON *et al.* 1988), high soluble solids concentration was correlated with low mass per fruit, and with high pH (Table 1). Also, low mass per fruit was correlated with high pH in the F_2, an association not previously detected. Correlations between traits based upon F_3 family means were weaker than those found in F_2, but indicated the same relationships. Solids were negatively correlated with fruit diameter in the Israel F_3, and positively correlated with pH in both F_3 trials.

Fruit diameter of the F_3-IS families was correlated with mass per fruit of the F_3-CA families and the F_2 plants. Unfortunately, fruit diameter and mass per fruit were not measured in the same environment, however it is not surprising that they would be closely associated, as fruit diameter resembles a cube-root transformation of mass per fruit.

Number of QTLs identified for each trait

Mass per fruit: A total of 11 QTLs influencing mass per fruit were detected (Table 2; Figure 4): in all cases, the CM allele reduced mass per fruit. Analysis of the F_2 revealed seven factors, on chromosomes *1, 2, 3, 6* (two), *7*, and *11*. Five of these regions showed effects on mass per fruit in the F_3-CA trial, as well as additional regions of chromosomes *4, 9,* and *12*. Another QTL affecting mass per fruit in F_3 mapped to chromosome *3*, just outside the likelihood interval of a F_2 QTL—this region was associated with segregation distortion, possibly reducing the accuracy of the results from F_3 progeny testing.

A total of four QTLs influencing fruit diameter were detected in the Israel F_3 (Table 2; Figure 4): in all cases, the CM allele reduced fruit diameter. Regions of chromosomes *3* and *11* associated with reduced fruit diameter were also associated with reduced mass per fruit. Regions of chromosomes 4 and 10 affected fruit diameter in F_3-IS, but showed no significant effect on mass per fruit in either F_2 or F_3-CA trials.

Soluble solids concentration: A total of seven QTLs influencing soluble solids concentration were detected (Table 3; Figure 4): in all cases, the CM allele elevated soluble solids concentration. Analysis of the F_2 revealed four factors, one each on chromosomes *2* and *3*, and two on chromosome *6*. At least two of these regions showed effects on soluble solids in the F_3-CA trial, and three in the F_3-IS trials. An additional QTL likelihood interval from the F_3-CA trial may encompass the two remaining F_2-QTLs (Figure 4, SS6c), while the F_3-IS trial revealed two new QTLs, on chromosomes *7* and *9*.

Fruit pH: A total of nine QTLs influencing fruit pH were detected (Table 4; Figure 4). Analysis of the F_2 revealed five factors, on chromosomes *4, 6, 7, 8,* and *10*. Three of these chromosomal regions (chromosomes *4, 7,* and *10*) showed effects on pH in the California F_3, as well as four additional regions on chromosomes *1, 3, 4,* and *8*. Only two QTLs affecting fruit pH were detected in the Israel F_3, in regions of chromosomes *4* and *6* which showed effects on pH in the F_2.

In six cases, QTL alleles from CM increased pH, while in the other three the CM alleles decreased pH. Although insufficient data is available on the pH of CM, pH of the E × CM progeny tended to be similar to that of E (Figure 1), suggesting that E and CM may themselves have similar pH. In a previous study, wild and domestic strains with similar pH were also found to segregate for alleles with compensatory effects (PATERSON *et al.* 1988). Recombination of such alleles represents a genetic basis for the phenomenon of *transgression*, the occurrence of progeny with phenotypes more extreme than either parent (SIMMONDS 1981).

Gene action of individual QTLs

The gene action of individual QTLs was evaluated by comparing the fits of individual QTL models constrained to "pure" additivity ($d = 0$) or "pure" dominance ($d = \pm a$). When it was possible to reject a purely additive model, for example, we say the locus *was consistent with* additivity. This is not meant to be interpreted as saying that the QTL exactly fits a purely additive model, an assertion which cannot be made from these data.

Mass per fruit: Most CM factors affecting mass per fruit were consistent with additivity (Table 2). In the F_2, both dominance ($d < 0$) and recessiveness ($d > 0$)

TABLE 2

Biometrical parameters of individual QTLs affecting mass per fruit (Mf) or fruit diameter (Fd), in three-environment trial of *L. esculentum* × *L. cheesmanii*

Locus	Trial	LOD	%Var	*a*	*d*	*d/a*	Mode
Mf1a	F_2	8.89	18.0	−0.148	−0.003	0.02	A
	CA	5.60	11.7	−0.109	0.014	−0.13	A
Mf2a	F_2	4.97	8.4	−0.087	−0.043	0.49	DA
	CA	2.87	5.0	−0.058	−0.040	0.69	DA
Mf3a	F2	21.51	42.0	−0.237	−0.002	0.01	A
Mf3b	CA	10.54	17.9	−0.134	0.029	−0.22	A
(*Fd*)	IS	4.00	20.8	−0.474	−0.405	0.85	D
Mf4a	CA	3.23	4.7	−0.062	−0.039	0.63	DA
Fd4a	IS	3.06	10.5	−0.170	0.131	−0.77	RA
Mf6a	F_2	6.88	13.2	−0.121	0.013	−0.11	A
	CA	6.53	11.5	−0.105	0.032	−0.30	A
Mf6b	F2	3.58	6.9	−0.081	0.018	−0.22	AR
Mf7a	F_2	5.14	10.9	−0.106	0.053	−0.50	A
	CA	3.21	6.7	−0.074	0.053	−0.72	RA
Mf9a	CA	2.76	6.7	−0.084	−0.002	0.02	ADR
Fd10a	IS	2.42	4.0	−0.126	−0.050	0.40	AD
Mf11a	F_2	5.18	11.4	−0.122	−0.008	0.07	A
	CA	2.45	5.9	−0.078	−0.005	0.06	A
(*Fd*)	IS	2.30	12.3	0.304	−0.316	−1.04	R
Mf12a	CA	3.81	5.9	−0.033	−0.090	2.73	D

Individual QTL loci (Locus) are named by trait (abbreviations indicated in titles), and chromosome number (*1–12*). In cases where multiple QTLs affecting a trait were found along the same chromosome, the QTLs are distinguished by letters indicating the temporal order in which the QTLs were discovered (*e.g.*; *Mf6a*, *Mf6b*). The environment(s) in which a QTL was detected are indicated (Trial: F_2 = F_2 trial in Davis, California; CA = F_3 trial in Gilroy, California; IS = F_3 trial in Rehovot, Israel). QTLs from different trials are listed under the same locus designation when likelihood intervals coincided closely (see Figure 4 for likelihood intervals). The LOD score (LOD) and percent phenotypic variance explained (% Var) by the QTL are presented, from single-locus models with unconstrained gene action. The additive effect (*a*, expressed as the effect of substituting a wild allele for a domestic allele), dominance deviation (*d*), and ratio of dominance to additivity (*d/a*) for each QTL are presented in original units for soluble solids concentration (°Brix) and pH, and in $\log_{10}$ (grams) for mass per fruit. The possible pure modes of gene action (Mode) for each QTL are indicated, based on testing of additive (A) and dominant (D, R) models as described in MATERIALS AND METHODS (*e.g.*, if $d = 0$ then A, if $d = a$ then D, if $d = -a$ then R). As shown in Figure 3, if a model reduced likelihood by 10-fold or more, it was deemed unlikely. When two pure modes of gene action could not be deemed unlikely, the more likely mode was listed first (*e.g.*, for *Mf2a*, dominance (D) was most likely but additivity (A) could not be deemed unlikely, thus the mode for this locus is denoted DA).

were deemed unlikely for five of the seven QTLs (MF1a, 3a, 6a, 7a and 11a), dominance alone was unlikely for one (MF6b), and recessiveness alone was unlikely for the other (MF2a). Of the four F_2 likelihood intervals which also showed effects in the F_3, three were consistent with the same type of gene action in both generations (MF2a, 6a, 11a), and for the fourth, recessiveness could no longer be falsified (MF7a).

Soluble solids concentration: Most CM factors affecting soluble solids concentration were consistent with additivity and/or dominance (Table 3). In the F_2, recessiveness was deemed unlikely for all of the four QTLs (SS2a, 3a, 6a, 6b), and dominance was unlikely for one QTL (SS3a). Of the three F_2 likelihood intervals which also showed effects in the F_3, one (SS2a) was consistent with the same type of gene action in both generations. The other two were consistent with similar modes of gene action in both generations: in one case (SS3a), additivity was most likely in both generations, but recessiveness could not be falsified in the F_3; and in the other case (SS6a), additivity was most likely in both generations, but dominance could not be falsified in the F_2.

Fruit pH: Most CM factors affecting fruit pH showed were consistent with additivity and/or recessiveness (Table 4). In the F_2, dominance was deemed unlikely for all but one QTL (pH6a), and even for this QTL, additivity was more likely than dominance. Of the four remaining F_2 QTLs, both additivity and dominance were deemed unlikely for one (pH7a), recessiveness was more likely than additivity for two (pH4a and pH10a), and additivity was more likely than recessiveness for one (pH8a). Each of the five F_2 likelihood intervals showed effects in at least one F_3 trial—two of these were consistent with the same type of gene action in both generations (pH4a, 6a). The other three showed similar modes in both generations: in one case (pH7a), additivity was unlikely in F_2 but not F_3, in the second case (pH10a), additivity was unlikely in F_3 but not F_2, and in the third case (pH8a), additivity was more likely than recessiveness in F_2 while recessiveness was more likely in F_3.

Epistasis between QTLs: Only minimal evidence of epistasis was found. In the F_2 generation, two-way interactions between unlinked markers were significant ($P < 0.05$) in about 8%, 7%, and 10% of cases, for mass per fruit, soluble solids concentration, and

TABLE 3

Biometrical parameters of individual QTLs affecting soluble solids concentration (SS), in three-environment trial of *L. esculentum* × *L. cheesmanii*

Locus	Trial	LOD	%Var	*a*	*d*	*d/a*	Mode
Ss2a	F_2	7.69	12.6	0.575	0.479	0.83	DA
	CA	3.43	8.2	0.307	0.185	0.60	DA
	IS	3.51	7.0	0.468	0.157	0.34	DA
Ss3a	F_2	11.97	27.5	1.176	−0.028	−0.02	A
	CA	3.38	6.0	0.326	−0.011	−0.03	AR
	IS	4.11	8.3	0.556	−0.061	−0.11	AR
Ss6a	F_2	3.89	7.3	0.563	−0.016	−0.03	AD
	IS	4.91	9.2	0.380	−0.083	−0.22	A
Ss6b	F_2	3.35	6.6	0.503	0.228	0.45	DA
Ss6c	CA	4.10	11.2	0.504	0.432	0.86	DA
Ss7a	IS	2.96	6.2	−0.344	0.392	−1.14	R
Ss9a	IS	4.30	12.0	0.650	−0.103	−0.16	AR

See Table 2 légend.

TABLE 4

Biometrical parameters of individual QTLs affecting fruit pH (pH), in three-environment trial of *L. esculentum* × *L. cheesmanii*

Locus	Trial	LOD	%Var	*a*	*d*	*d/a*	Mode
pH1a	CA	3.80	5.5	−0.036	−0.014	0.39	AD
pH3a	CA	3.52	6.9	−0.040	−0.014	−0.35	RA
pH4a	F_2	2.49	5.1	−0.049	0.039	−0.80	RA
	CA	2.65	4.2	−0.030	0.021	−0.70	RA
	IS	6.05	13.9	−0.109	0.035	−0.32	AR
pH4b	CA	3.57	5.2	0.029	0.027	0.93	D
pH6a	F_2	3.87	7.3	0.065	0.007	0.11	AD
	IS	3.53	8.5	0.082	−0.001	−0.01	AD
pH7a	F_2	8.85	15.2	0.075	−0.074	−0.99	R
	CA	3.20	4.8	0.034	−0.012	−0.35	AR
pH8a	F_2	2.54	8.8	0.072	−0.022	−0.31	AR
	CA	4.12	28.3	0.076	−0.055	−0.72	RA
pH10a	F_2	2.79	4.7	−0.053	0.035	−0.66	RA
	CA	3.59	7.9	−0.035	0.049	−1.40	R
pH12a	CA	2.41	4.4	−0.029	0.027	−0.93	RA

See Table 2 legend.

fruit pH, respectively—only slightly more frequently than would be expected by chance.

The strongest evidence of epistasis that was found involved a CM factor near *TG35* on chromosome *9*, which appears to interact with CM factors on chromosomes *2, 3, 4, 5* and *7*, in elevating soluble solids concentration (data not shown). In three cases, additive * additive interaction explained most of the effect, the other cases being large additive * dominant and dominant * additive. None of the six regions involved showed a significant effect on soluble solids concentration individually, although the *TG35* region itself was only slightly below significance.

Pleiotropic effects of QTLs: QTLs affecting different traits fell near one another more frequently than would be expected by chance. This suggests that the observed correlations between traits (Table 1), particularly the strong correlation between mass per fruit and soluble solids concentration, may be partly due to pleiotropic effects of single QTLs. In the E × CM F_2, 1-LOD (~90%) likelihood intervals for QTLs affecting mass per fruit, soluble solids concentration, and fruit pH span 16%, 11%, and 11% of the genome, respectively. On a random basis, likelihood intervals for QTLs affecting mass per fruit and soluble solids concentration should coincide over (16%) * (11%) = ~2% of the genome, or about one likelihood interval. In fact, likelihood intervals for three of the six QTLs affecting mass per fruit coincide closely with those of QTLs affecting soluble solids concentration. This suggests that either some individual QTLs have pleiotropic effects (GRUNEBERG 1938) on both soluble solids and mass per fruit, or that different QTLs affecting these traits tend to be clustered together into closely linked groups. Pleiotropy or close linkage between mass per fruit and soluble solids concentration has been suggested by numerous other studies, using both classical analyses (GOLDENBERG and PAHLEN 1966; IBARBIA and LAMBETH 1969; RICK 1974) and QTL mapping (PATERSON *et al.* 1988, 1990). Associ-

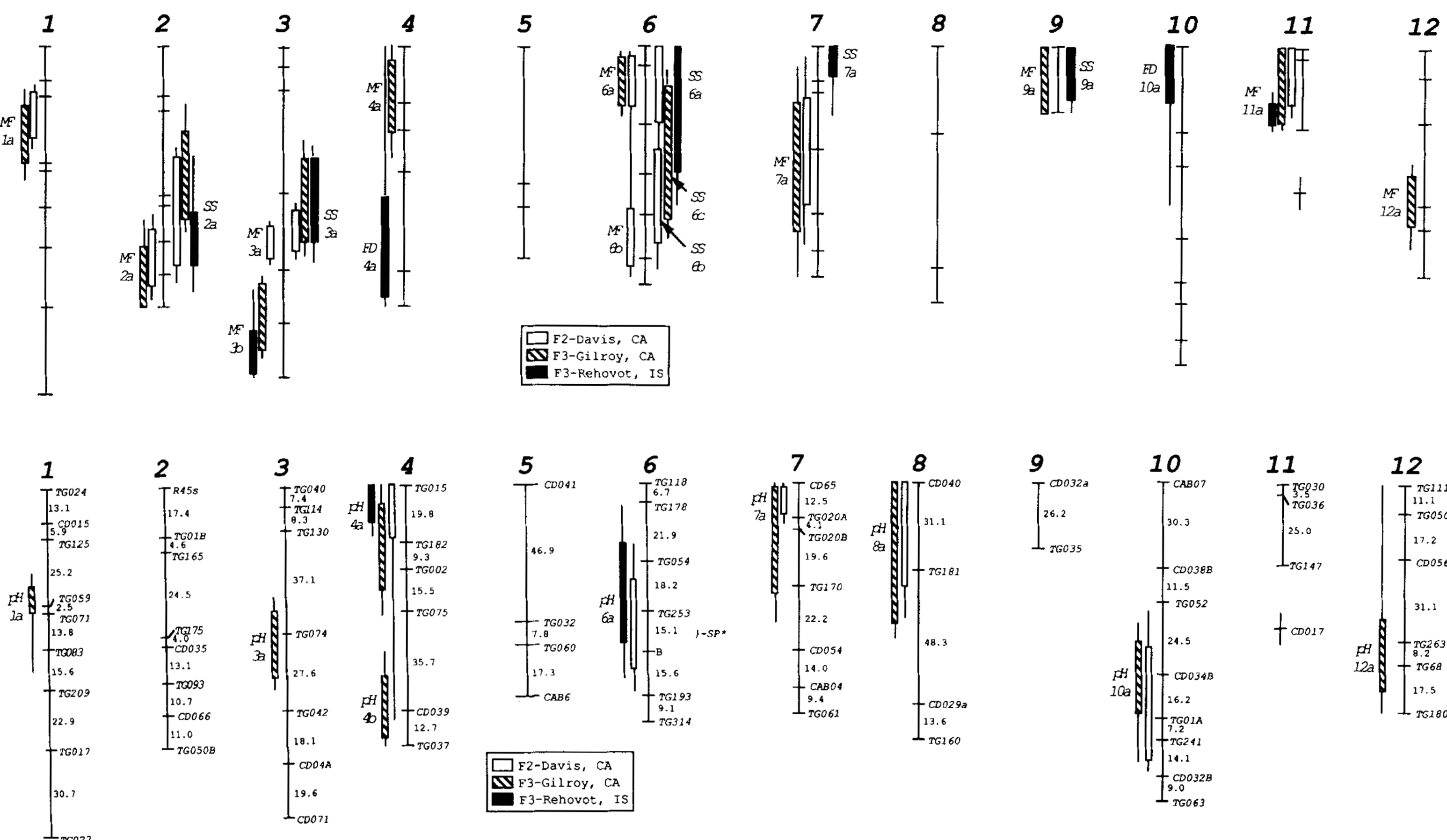

FIGURE 4.—Likelihood intervals for QTLs mapped in F_2, F_3-California, and F_3-Israel trials. The position of a QTL is shown as the interval over which the LOD score is within 1 or 2 log-units of that at the most likely position, indicating the region over which the model's probability of giving rise to the data is at most 10- or 100-fold less than at the most likely position. Bars indicate 1-LOD (10-fold) likelihood intervals, with whiskers (lines extending beyond bars) indicating 2-LOD (100-fold) likelihood intervals. Individual QTLs have been named according to trait (MF = mass per fruit; FD = fruit diameter; SS = soluble solids concentration; pH = pH), chromosome (#, 1 to 12), and then sequentially by order of discovery (a, b, etc.). Using this nomenclature, QTL locations have been cross-referenced with phenotypic effects and types of gene action presented in Tables 2–4. (Top) Mass per fruit or fruit diameter (left side of chromosome), and soluble solids concentration (right side of chromosome). (Bottom) Fruit pH, with marker names and map distances indicated.

TABLE 5

Comparison of F_2 phenotype and F_2-QTLs, as predictors of F_3 progeny phenotypes

		Percent phenotypic variance explained by[a]			
Trait	Narrow-sense heritability	F2-QTL model	Residual factors[b]	Classical F_3/F_2 regression	QTL-based F_3/F_2-expected regression
Mass per fruit	0.45	72	5	45	30
Fruit pH	0.25	34	13	14	15
Soluble solids conc.	0.15	44	11	5	11

[a] All variances were significantly different from zero, at $P \leq 0.01$.
[b] Including environmental effects and measurement error, estimated from variation among replicates of the *L. esculentum* parent.

ations between mapped QTLs affecting the other possible combinations of traits (*e.g.*, mass per fruit × pH, soluble solids × pH) are no greater than would be expected by chance—only one of the four QTLs affecting pH overlaps with a QTL affecting mass per fruit, and another with a QTL affecting soluble solids concentration.

One of the chromosomal regions affecting both mass per fruit and soluble solids concentration is on chromosome *6*, near the *B* locus. This corresponds closely to the region of the *sp* locus, which has previously been suggested to affect mass per fruit and soluble solids concentration in other tomato populations (EMERY and MUNGER 1966; PATERSON *et al.* 1988). The *sp* gene did segregate in the present population, and was scored—however, multipoint linkage analysis suggested that, for several individuals, genotype at flanking RFLP markers was inconsistent with genotype inferred from the *sp* phenotype (termination of apical elongation by formation of a flow cluster). This may indicate that other genes, segregating independently from the locus on chromosome *6*, can influence the *sp* phenotype to some degree.

Predictive value of multiple-QTL models

Phenotypic and genetic variance explained: After identifying individual QTLs by single-locus models, we combined the QTLs into a multi-locus model to determine how much of the phenotypic variance among the F_2 progeny was explained (Table 5). The seven QTLs affecting mass per fruit explain 72% of the phenotypic variance, the four QTLs affecting soluble solids concentration explain 44% and the five QTLs affecting pH explain 34%. Environment plus measurement error were estimated to contribute about 5%, 11% and 13% to phenotypic variation in mass per fruit, soluble solids concentration, and fruit pH, respectively, based on variation among the E check plants grown with the F_2 generation (Figure 1, insets). Eliminating this source of variance, the identified QTLs account for 76%, 49%, and 39% of the genetic variance in mass per fruit, soluble solids concentration and pH, respectively. We emphasize that phenotypic variance explained by our model probably underestimates total genetic variance, because only additive, dominant, and recessive genetic components are included in the model. Additional genetic variance may be due to QTLs with effects too small to detect in this experiment, or epistasis (which also was generally too small to detect in this experiment; see below).

The variance explained by QTLs identified in the F_2 indicates that the phenotypic variation in these traits has a large genetic component. However, these numbers (72%, 44% and 34%) cannot be interpreted as a measure of resemblance between relatives ("heritability"), as they are biased upward by both nonadditive genetic factors, and genotype × environment interaction. Direct estimates of heritability, from resemblance between F_2 individuals and their F_3 progeny, are presented below.

Prediction of progeny phenotypes: Having determined both phenotype and genotype of the F_2 individuals, as well as phenotype of the F_3 selfed progeny of the F_2, it was possible to compare the predictive value of classical phenotypic data with that of QTL mapping. This was done using data from the F_2 and Gilroy, California, F_3 trials, as follows: (1) The predictive value of F_2 phenotypes was assessed, by calculating the regression of F_3 progeny phenotype on F_2 parent phenotype, for each measured trait. In addition, an estimate of the heritability of each trait was calculated from the parent-progeny regression (SMITH and KINMAN 1965), using standard units (FREY and HORNER 1957). (2) The predictive value of F_2-QTL genotypes was assessed, by calculating the regression of F_3 progeny phenotype on an "F_2-expected phenotype." This "F_2-expected phenotype" was computed for each F_2 individual, from the repertoire of QTLs the individual was inferred to carry (this is readily done, since we know the locations and phenotypic effects of the QTLs from mapping, and can infer the "QTL genotype" of an individual from its genotype at flanking RFLPs). If our models included all existing QTLs (*i.e.*, not just those we can detect at the resolution of this experiment) and were perfect in determining location and phenotypic effect, the "F_2-expected phenotype" for an individual would be equal to its ob-

served phenotype minus environmental noise (and other nongenetic factors).

The predictive value of the F_2 QTLs was inversely related to the heritability of the traits studied (Table 5). For the trait of highest heritability, mass per fruit, F_2 progeny phenotypes explained 44% of the variance among F_3 family means, while "F_2-expected phenotype" explained only 30%. For the trait of intermediate heritability, fruit pH, F_2 progeny phenotypes explained 14% of the variance among F_3 family means, and "F_2-expected phenotype" explained 15.4%, a similar amount. For the trait of lowest heritability, soluble solids, F_2 progeny phenotype explained only 4.7% of the variance among F_3 family means, while "F_2-expected phenotype" explained 11%, more than twice as much.

Effects of environment on QTL expression: By growing populations in three different environments (Davis, California, 1987; Gilroy, California, 1988; and Rehovot, Israel, 1988), it was possible to assess the influence of genotype × environment interaction, as reflected by the expression of individual QTLs. Among a total of 29 putative QTLs mapped (Tables 2–4), only 4 (14%) were detected in all three environments, 10 (34%) were detected in two environments, and 15 (52%) were detected only in a single environment.

The two California environments appeared to be more similar to one another than either was to the Israel environment. Across the three traits studied, the F_2 and F_3-California trials shared 11/25 QTLs (44%), the F_2 and F_3-Israel trials shared 7/20 QTLs (35%), and finally, the F_3-California and F_3-Israel trials shared 5/26 QTLs (19%) (summarized from Tables 2–4). We note, however, that the comparison across environments is confounded by two factors: (1) generation (F_2 *vs.* F_3), and (2) measurement of fruit diameter rather than mass per fruit in Israel.

Similarity in location of CM and CL QTLs: Many QTLs mapped in the E × CM F_2 are located near QTLs affecting the same traits in the E × *L. chmielewskii* (CL) backcross (Paterson *et al.* 1988). The 1-LOD (~90%) likelihood intervals for QTLs from CL affecting mass per fruit, soluble solids concentration and fruit pH spanned 15%, 10%, and 14% of the genome (similarly to those of CM—see Pleiotropy section, above). Likelihood intervals in the E × CM F_2 and E × CL BC1 coincide closely for 3 of 7 QTLs (43%) affecting mass per fruit, 2 of 4 QTLs (50%) affecting soluble solids concentration, and 3 of 5 QTLs (60%) affecting pH. [Figure 5; also, note that the reciprocal values for CL are 3/6 (50%), 2/4 (50%), and 3/5 (60%).] Across the three traits, 50% (8 of a possible 16) of likelihood intervals coincide, more than three times the ~15% which would be expected by chance (see Pleiotropy section, above). This estimate

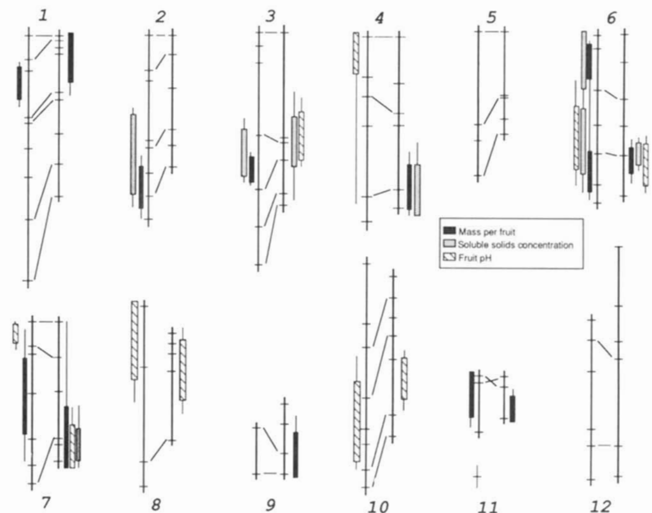

FIGURE 5.—QTL likelihood maps of E × CM F2 (left) and E × CL BC1 (right) chromosomes. RFLP markers which were common to the two maps are indicated by diagonal lines. The position of a QTL is shown as the interval over which the LOD score is within 1 or 2 log-units of that at the most likely position, indicating the region over which the model's probability of giving rise to the data is at most 10- or 100-fold less than at the most likely position. Bars indicate 1-LOD (10-fold) likelihood intervals, with whiskers (lines extending beyond bars) indicating 2-LOD (100-fold) likelihood intervals. For data analysis, 1-LOD likelihood intervals of CM and CL were considered to overlap if they fell largely within the same map interval. Of the 16 likelihood intervals for CM and 15 for CL, the following 8 were considered to overlap: (1) chromosome *1* mass per fruit, (2) chromosome *3* soluble solids concentration, (3–5) chromosome *6* mass per fruit, soluble solids concentration, and pH at lower end of chromosome, (6) chromosome *8* pH, (7) chromosome *10* pH, (8) chromosomal *11* mass per fruit. In addition, likelihood intervals affecting mass per fruit on chromosome *7* were close to one another, however we considered the overlap insufficient to call a match.

of similarity between CM and CL may be conservative, in that E × CL was a backcross, thus any recessive factors from CL would not have been detected. It is clear, however, that the extreme phenotypes of CM and CL (relative to E) appear to be associated with genetic factors in some common regions of the genome.

DISCUSSION

QTL mapping in F_2 populations: In the F_2 generation of E × CM, we were able to map seven putative QTLs affecting mass per fruit, four affecting soluble solids concentration, and five affecting fruit pH. These QTLs have been described in much the same manner as "major genes"; by chromosomal location, by magnitude of effect on phenotype, by additive effects and dominance deviations, by interaction (or lack thereof) with unlinked genetic factors, and by multiple effects (or lack thereof) on different traits. In addition, QTLs may be described by their sensitivity to environmental factors—a few QTLs were detected in all three environments studied, while most

were only detected in one of the three environments. Consequently, studies done in a single environment are likely to underestimate the number of QTLs which can influence a trait.

In total, the QTLs which could be mapped in the F_2 accounted for 72% of the phenotypic variance in mass per fruit, 44% in soluble solids concentration, and 34% in fruit pH. It was not surprising that we were able to explain so much more variation in mass per fruit than in the other two traits, since the parents differed by 20 SD in mass per fruit, but only 10 SD in soluble solids concentration and even less in pH. The remaining variation, which could not be explained by the QTL models, comes from at least four sources: (1) environment plus measurement error—based on variation in phenotype of the homozygous E parent, these accounted for 5% of the phenotypic variance in mass per fruit, 11% in soluble solids concentration, and 13% in fruit pH. After adjusting for these nongenetic factors, the mapped QTLs account for 76%, 49%, and 39% of the genetic variance in the three traits, respectively–suggesting the presence of unmapped genetic effects. These could include (2) QTLs with effects too small to detect with confidence in this experiment, (3) interactions between QTLs, which were generally too small to detect in this experiment but could still contribute to phenotypic variance, and (4) interactions of individual F_2 genotypes with environmental variation within the F_2 experiment (*e.g.*, environmental variation across the field, or in statistical terms, block × treatment interaction). More generally, we emphasize that the percentage of phenotypic variance explained by genetic factors cannot be used as an indication of the "quality" of an experiment, without also considering variance due to nongenetic factors (environment and measurement error). One might more directly express the information available from a QTL mapping experiment by determining the percentage of genetic variance explained by QTLs.

QTLs explaining small portions of the phenotypic variance far outnumber those explaining larger portions (Figure 6), with many explaining 4–8% but only one explaining more than 40% (the smallest effect which could be detected in this experiment was about 4% of variance explained). These results, along with many earlier studies, support a model for quantitative inheritance wherein effects of individual factors range from essentially qualitative down to vanishingly small (see LANDE and THOMPSON 1990 for a more detailed discussion of this, and also many relevant citations). Prior results have shown that as QTLs with large effects are fixed in a population, it becomes possible to detect factors with progressively smaller effects (SHRIMPTON and ROBERTSON 1988; PATERSON *et al.* 1990). As the variance explained by a QTL decreases,

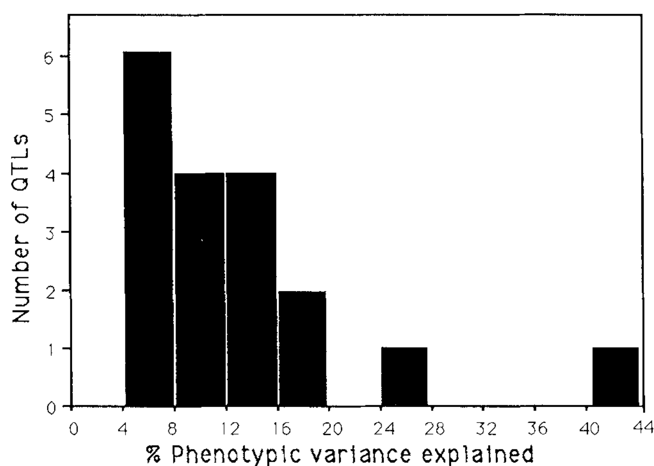

FIGURE 6.—Distribution of phenotypic variance explained by individual QTLs mapped in the E × CM F_2 population. QTLs affecting mass per fruit, soluble solids, and pH are presented in common units (percent phenotypic variance explained). Individual QTLs were found to explain from as much as 42% of the phenotypic variance in a trait, down to as little as 4.7% (near the significance threshold). Because the significance threshold for QTLs was about 4% of variance explained in this experiment, no data is available on the 0–4% interval.

the number of progeny which must be studied in order to detect the QTL increases (see LANDER and BOTSTEIN 1989). Only QTLs with sufficiently large effects will be detected in a particular cross, while those with smaller effects will go unnoticed. Consequently, estimates of QTL numbers should be considered as lower bounds.

Gene action of individual QTLs: In a F_2 population, one can determine the effect of different gene dosages on phenotype (*e.g.*, gene action), because all three possible gene dosages at a locus are represented. This cannot be done in backcross populations, which lack one parental dosage, or recombinant inbred populations, which lack the heterozygous dosage. For the same reason, a F_2 population can be used to map recessive factors from *either* parent, unlike a backcross.

For most QTLs studied herein, at least one of the three types of gene action tested (additive, dominant [$d > 0$], recessive [$d < 0$]), could be deemed unlikely by 1 LOD (10-fold) or more (Tables 2–4). In only a few cases were two modes of inheritance found unlikely, suggesting that the corresponding QTLs were clearly additive, dominant or recessive. The inability to reject modes of inheritance is limited partly by the number of individuals examined, but may also indicate QTLs with "partially dominant" (or recessive) gene action.

Gene action of a QTL is most accurately estimated from the phenotype of F_2 individuals with known RFLP genotypes. However, it can also be estimated by determining the average phenotype of selfed or intercrossed progeny, from F_2 individuals of known genotype. This "progeny testing" approach requires

no additional RFLP analysis, and permits one to study many progeny in replicated trials. Unfortunately, progeny testing provides successively less accuracy at estimating gene action with each generation of selfing (or intercrossing). For example, modes of inheritance deemed unlikely based on phenotypes of F_2 individuals could often not be rejected based on average phenotypes of F_3 families (Tables 2–4, also see *Gene action*, in RESULTS). This is consistent with theoretical expectations. The ability to detect a dominance deviation is reduced by half with each generation of selfing (intercrossing), as heterozygosity is lost: for example, a heterozygote has expected phenotype $a + d$, but its progeny have expected phenotype $a + (1/2)d$. After several generations of selfing, one creates a largely homozygous population (such as recombinant inbreds), and can no longer distinguish additivity from dominance.

And what of epistasis? We found only a little evidence of two-way interaction between unlinked genetic factors, however our power to detect even simple epistatic effects such as these was limited (in view of the large number of potential pairwise interactions and the relatively small sampling of any particular combination of genotypes at a pair of loci). Similarly, prior marker-based studies (PATERSON *et al.* 1988, 1990), including one study using much larger populations (EDWARDS, STUBER, AND WENDEL 1987), found little evidence of epistasis. However, a considerable body of research in quantitative genetics suggests that interactions between QTL alleles at different loci have a considerable influence on phenotype (SPICKETT and THODAY 1966; also see ALLARD 1988). Further, despite the apparent lack of epistasis among QTLs with relatively large effects, QTLs with smaller effects have been shown to function differently in different genetic backgrounds (TANKSLEY and HEWITT 1988, A. H. PATERSON and S. D. TANKSLEY, unpublished results). Perhaps some QTLs *appear* to have small effects because they are dependent upon interaction with other loci, and in small populations the optimal allele configurations occur only rarely. In other words, studies to date may have *preferentially* identified QTLs which function independently of unlinked genetic factors. The importance of epistasis in quantitative inheritance may be better elucidated in the future, by studying larger populations, more closely-spaced markers, and/or specially constructed genetic stocks carrying particular combinations of QTLs (SPICKETT and THODAY 1966), preferably in an otherwise homogeneous background. Such studies may be necessary to reveal the genetic basis of subtle phenotypic differences which distinguish successful crop varieties from their lesser brethren.

A relatively small role of epistasis would bode well for extraction of agriculturally useful traits from wild species, as a single factor is much easier to identify and extract than a pair (or more) of interdependent factors. This becomes especially important if desirable QTLs are linked to undesirable genes from a wild source (such as reduced mass per fruit and high soluble solids concentration from CM), because it is easier to break such linkages at one locus than at two or more.

Resolution of QTL locations: The ability to localize any given QTL is limited by the number of meioses studied, and by noise introduced from environmental effects, measurement error, segregation of other QTLs, and interaction of QTLs with each other and/or environment. F_2 populations are more informative than backcross populations of the same size, since twice as many meioses are studied in each individual. This helped us to discern multiple QTLs on a chromosome, affecting the same trait (Figure 3, chromosome *6*; two QTLs each for mass per fruit and soluble solids concentration). One might gain similar information using a recombinant inbred population (BURR *et al.* 1988), but would sacrifice the ability to distinguish additive factors from dominant factors (without making additional crosses). However, two-generation crosses (such as F_2, backcross or recombinant inbred) provide only approximate localization of QTLs. More accurate localization can be obtained by performing additional backcrosses to produce isogenic lines differing only in the region containing a QTL; this will eliminate the majority of the genetic variance and will make it possible to dissect the remaining interval by examining various recombinants for flanking markers (*cf.* PATERSON *et al.* 1990).

Breeding value of F_2 QTL maps: In our experiments, the value of QTL mapping for predicting progeny phenotype was correlated inversely with the heritability of the trait, making a greater contribution to predicting traits of low heritability (Table 5—also see RESULTS). Although additional experiments are needed to assess the strength of this correlation, it does seem intuitive—for a highly heritable trait, by definition, an individual's phenotype is a good indicator of its "breeding value" (potential for producing desirable progeny). For a less heritable trait, an individual's phenotype is more greatly influenced by nongenetic factors such as environment. For such traits, phenotype is thus is a less effective indicator of an individual's breeding value, and QTL genotype may be relatively more informative. For a mapped QTL, the phenotypic effect is estimated from the data on many individuals, thus the influence of nongenetic factors should be reduced.

The observation that mapped QTLs (*e.g.*, the "F_2-expected phenotype") added more information about traits of low heritability than traits of high heritability supports prior predictions and theoretical expecta-

tions. Several authors have predicted that QTL mapping would prove especially useful in breeding for traits influenced greatly by environment (BURR *et al.* 1983; STUBER and EDWARDS 1986; SOLLER and BECKMANN 1988). Further, quantitative genetic theory indicates that marker-assisted selection should yield a greater relative improvement in selection efficiency for low heritability traits than for high-heritability traits (LANDE and THOMPSON 1990). However, we suggest that additional experiments are needed to verify the strength of the relationship between heritability and QTL-based predictions, studying a larger assortment of traits in species with different breeding systems (*e.g.* polyploids and outcrossing species, in addition to the diploid self-pollinated species studied here).

To efficiently breed for traits of high heritability, the logical first step might be classical phenotypic selection to fix QTLs of high heritability, rather than a QTL mapping experiment. The possible locations of these QTLs might be determined after several generations of selection, by simply comparing the genotype of the selected stocks to that of the recipient stock (TANKSLEY and HEWITT 1988). Once QTLs of large effect have been fixed, however, the heritability of the trait is reduced (genetic variance is reduced, while nongenetic variance is constant), creating a situation where QTL mapping might then offer additional gains beyond those readily achievable by classical means. We note that even for QTLs of high heritability, knowing map position (relative to linked markers) might permit one to more rapidly introgress QTLs into a new background, and would permit one to do so in an environment where the trait cannot be assayed. Finally, QTL mapping clearly provides information which might ultimately be useful in cloning QTLs (PATERSON *et al.* 1990a).

Sensitivity of QTLs to environment: Individual QTLs appear to show a range of sensitivities to environment, as some QTLs were detected in all three test environments while many could be detected only in a single environment. This is not particularly surprising. Classical plant breeders routinely find that genotypes which perform well in one environment are not well-suited to other environments. Sometimes these differences among genotypes can be attributed to relatively simply inherited attributes such as susceptibility to particular strains of pest, photoperiod response, or vernalization requirement. However, differences in adaptation of plant or animal genotypes may also be due to environment-sensitive QTLs.

QTLs which function consistently over a range of environments are preferred for breeding, however the additional use of environment-specific QTLs may further improve agricultural productivity. By constructing near-isogenic stocks and testing individual QTLs as one might test advanced breeding lines (A. H. PATERSON, in preparation) it might be possible to define production environments under which individual QTLs function reliably. Further, by combining several QTLs with different environment specificities into a single genotype, one might elicit an improvement in phenotype which is somewhat buffered against the vagaries of environment. In fact, this may be what is accomplished in classical plant breeding by making selections in the target environment then testing in a number of different environments. However, use of genetic markers to identify and manipulate the genes of interest might greatly accelerate the process.

Similar locations of QTLs in different species; evidence for variation in orthologous genes? Although CM and CL are only distantly related (RICK 1979, MILLER and TANKSLEY 1990), we previously noted that they resemble each other and differ from E in several ways. In particular, both CM and CL show much lower mass per fruit, much higher soluble solids concentration, and somewhat higher pH than E. Comparing the crosses E × CM and E × CL, many of the QTLs affecting a trait mapped to similar locations in the genome. There may be artifactual explanations of this, such as recombination suppression in some common regions of the genome making genes appear closer together recombinationally than they actually are physically. However, an intriguing possible explanation of this similarity is that phenotypic differences between E and CM, and between E and CL, may be due to allelic variation in some of the same genes. (Since the genes are in different species, rather than the "same" genes it is technically more accurate to refer to *orthologous* genes, derived from a common ancestral gene.)

The mapped QTLs almost certainly do not comprise the entire set of genes which affect the trait(s) under study. A large number of genes, encoding proteins involved in transcription, translation, energy metabolism, vegetative and reproductive development, senescence, and other fundamental processes of living organisms, could potentially influence the phenotypes we have studied. The mapped QTLs comprise only a subset of genes influencing the traits–specifically, the subset which shows allelic variation with a phenotypic effect which is detectable in the crosses and environments studied. The intriguing possibility is that a limited subset of the genes affecting a trait may account for much of the variation upon which selection acts, when a trait evolves. These genes may be the ones in which the organism can most easily tolerate mutations with a phenotypic effect.

If the similarity in QTL locations between these two distantly related species does represent allelic variation in orthologous genes, then one might contemplate "comparative QTL mapping," using infor-

mation from one population to predict properties of other related populations. The hypothesis that quantitative variation in phenotype is largely due to allelic variation in a limited number of loci has many important implications. Phenotypic variation within cultivated germplasm pools might be accounted for by a relatively small set of loci, and the map positions of many of these loci might be determined in a few well-chosen crosses between diverse cultivars. Germplasm collections might be treated as collections of variant alleles at a number of QTLs. Mapping QTLs in a small set of core genotypes might provide information relevant to much larger germplasm pools within a species or genus.

It is not inconceivable that some of the same (*e.g.*, orthologous) genes accounting for variation in a trait (for example, height or daylength sensitivity) could show polymorphism in different species, such as *Lycopersicon*, *Zea* or *Arabidopsis*. If so, it might be possible to clone QTLs for a trait of interest from relatively tractable species (*e.g.*, *Arabidopsis*), and then study the corresponding genes in different species. Of course, even if an orthologous gene is present in a different species, and even functions in the expression of the same trait, the gene will not necessarily be among those which account for *variation* in the trait (in the new species). Orthologous genes in different species may exhibit different levels of polymorphism, due to selection for primary, pleiotropic, or epistatic effects, evolutionary history (genetic drift or genetic bottlenecks), or other reasons. It has yet to be tested whether the sets of genes which show allelic variation influencing a quantitative trait are similar in *Lycopersicon*, *Zea*, and/or *Arabidopsis*.

QTL mapping may permit the study of genetic variation at "rate-limiting" steps in lengthy and complex "developmental pathways." This information may have practical utility in breeding agriculturally superior plants and animals, and basic utility in studying the biology of complex agricultural, physiological, and behavioral traits. Finally, by attaching quantitative information about phenotype to defined sites in the DNA of an organism, one establishes an effective medium for communicating information between applied life sciences and basic molecular biology, greatly expanding the repertoire of tools which might be brought to bear on crucial questions in agriculture, evolution, and medicine.

We thank MITZI AGUIRRE, MICHAEL BRIGGS, MARIA CARPENTER, JANICE CHEN, GERALD DICKINSON and SUSAN GREGORY for technical assistance. We thank the anonymous reviewers for thorough readings and valuable comments. We gratefully acknowledge support from U.S. Department of Agriculture/Competitive Research grant 88-37262-3921 to S.D.T., U.S.-Israel Binational Agricultural Research and Development Fund grant US-1388-87 to S.D.T. and D.Z., and System Development Foundation grant (G612), a National Science Foundation Grant, and a Markey Foundation grant to E.S.L.

LITERATURE CITED

ALLARD, R. W., 1956 Formulas and tables to facilitate the calculation of recombination values in heredity. Hilgardia **24:** 235–278.

ALLARD, R. W., 1988 Genetic changes associated with the evolution of adaptedness in cultivated plants and their wild progenitors. J. Hered. **79:** 225–238.

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

BOTSTEIN, D., R. L. WHITE, M. SKOLNICK and R. W. DAVIS, 1980 Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. **32:** 314–331.

BURR, B., F. A. BURR, K. H. THOMPSON, M. C. ALBERTSEN and C. W. STUBER, 1988 Gene mapping with recombinant inbreds in maize. Genetics **118:** 519–526.

BURR, B., S. V. EVOLA, F. A. BURR and J. S. BECKMANN, 1983 The application of restriction fragment length polymorphism to plant breeding, pp. 45–59 in *Genetic Engineering*, vol. 5, edited by J. K. SETLOW and A. HOLLAENDER. Plenum Press, N.Y.

DAVISSON, M. T., and AND T. H. RODERICK, 1981 Recombination percentages, pp. 416–427 in *Genetic Variants and Strains of the Laboratory Mouse*, edited by M. C. GREEN. Gustav Fisher Verlag, Stuttgart.

DONIS-KELLER, H., P. GREEN, C. HELMS, S. CARTINHOUR, B. WEIFFENBACH, K. STEPHENS, T. P. KEITH, D. W. BOWDEN, D. R. SMITH, E. S. LANDER, D. BOTSTEIN, G. AKOTS, K. S. REDIKER, T. GRAVIUS, V. A. BROWN, M. B. RISING, C. PARKER, J. A. POWERS, D. E. WATT, E. R. KAUFFMAN, A. BRICKER, P. PHIPPS, H. MULLER-KAHLE, T. R. FULTON, S. NG, J. W. SCHUMM, J. C. BRAMAN, R. G. KNOWLTON, D. F. BARKER, S. M. CROOKS, S. E. LINCOLN, M. J. DALY and AND J. ABRAHAMSON, 1987 A genetic linkage map of the human genome. Cell **51:** 319–337.

EAST, E. M., 1915 Studies on size inheritance in Nicotiana. *Genetics* **1:** 164–176.

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.

EMERY, G. C., and H. M. MUNGER, 1966 Effect of inherited differences in growth habit on fruit size and soluble solids in tomato. J. Am. Soc. Hortic. Sci. **95:** 410–412.

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

FREY, K. J., and T. HORNER, 1957 Heritability in standard units. Agron. J. **49:** 59–62.

GADISH, I., and D. ZAMIR, 1987 Differential zygotic abortion in an interspecific *Lycopersicon* cross. Genome **29:** 156–159.

GARVEY, T. C., and J. D. HEWITT, 1984 A survey of *Lycopersicon cheesmanii* for high soluble solids. Tomato Genet. Coop. Rept. **34:** 4–5.

GOLDENBERG, J. B., and A. PAHLEN, 1966 Genetic and phenotypic correlation between weight and dry matter content of tomato fruits and their heritabilities. Bol. Genet. **2:** 1–15.

GRUNEBERG, H., 1938 An analysis of the "pleiotropic" effects of a new lethal mutation in the rat (*Mus norvegicus*). Proc. R. Soc. Lond. B **125:** 123–144.

IBARBIA, E. A., AND V. N. LAMBETH, 1969 Inheritance of soluble solids in a large/small-fruited tomato cross. J. Am. Soc. Hortic. Sci. **94:** 496–498.

JOHANSSEN, W., 1909 *Elemente der exakten Erblichkeitsllehre.* Fischer, Jena.

KOSAMBI, D. D., 1994 The estimation of map distances from recombination values. Ann. Eugen. **12:** 172–175.

LANDE, R, 1981 The minimum number of genes contributing to

quantitative variation between and within populations. Genetics **99:** 541–553.

LANDE, R., AND R. THOMPSON, 1990 Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics **124:** 743–756.

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.

MATHER, K., and J. L. JINKS, 1971 *Biometrical Genetics*. Cornell University Press, Ithaca, N.Y.

MAEDA, T., 1939 Chiasma studies in the silkworm, *Bombyx mori L*. Jpn. J. Genet. **15:** 118–127.

MILLER, J. D., and S. D. TANKSLEY, 1990 RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. Theor. Appl. Genet. (in press).

MORGAN, T. H., 1914 No crossing over in the male of *Drosophila* of genes in the second and third pair of chromosomes. Biol. Bull. **26:** 195–204.

NILSSON-EHLE, H., 1909 *Kreuzunguntersuchungen an Hafer und Weizen*. Lunds. Univ. Aarskr. NF **5:** 1–122.

OSBORN, T. C., D. C. ALEXANDER and J. F. FOBES, 1987 Identification of restriction fragment length polymorphisms linked to genes controlling soluble solids content in tomato fruit. Theor. Appl. Genet. **73:** 350–356.

OTT, J., 1985 *Analysis of Human Genetic Linkage*. Johns Hopkins University Press, Baltimore, M.D.

PATERSON, A. H., E. S. LANDER, J. D. HEWITT, S. PETERSON, S. E. LINCOLN AND S. D. TANKSLEY, 1988 Resolution of quantitative traits into Mendelian factors, using a complete linkage map of restriction fragment length polymorphisms. Nature **335:** 721–726.

PATERSON, A. H., J. W. DEVERNA, B. LANINI and S. D. TANKSLEY, 1990 Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes, in an interspecies cross of tomato. Genetics **124:** 735–742.

RENWICK, J. H., and J. SCHULZE, 1965 Male and female recombination fraction of the nail patella:ABO linkage in man. Ann. Hum. Genet. **28:** 379–392.

RHOADES, M. M., 1941 Different rates of crossing over in male and female gametes of maize. J. Am. Soc. Agron. **33:** 603–615.

RICK, C. M., 1974 High soluble-solids content in large-fruited tomato lines derived from a wild green-fruited species. Hilgardia **42:** 493–510.

RICK, C. M., 1979 Biosystematic studies in *Lycopersicon* and closely related species of *Solanum*, pp. 677–678 in *The Biology of the Solonaceae*, edited by J. C. HAWKES, R. N. LESTER AND A. D. SKELDING. Academic Press, New York.

ROBERTSON, D. S., 1984 Different frequency in the recovery of crossover products from male and female gametes of plants hypoploid for B-A translocations in maize. Genetics **107:** 117–130.

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

SAX, K., 1923 The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics **8:** 552–556.

SHRIMPTON, A. E., and A. ROBERTSON, 1988 The isolation of polygenic factors controlling bristle score in *Drosophila melanogaster*. II. Distribution of third chromosome bristle effects within chromosome sections. Genetics **118:** 445–459.

SIMMONDS, N. W., 1981 *Principles of Crop Improvement*. Longman, N.Y.

SMITH, J. D., and M. L. KINMAN. 1965 The use of parent-offspring regression as an estimator of heritability. Crop Sci., **5:** 595–596.

SOLLER, M., and J. BECKMANN, 1988 Genomic genetics and the utilization for breeding purposes of genetic variation between populations, pp. 161–188 in *Proceedings of the 2nd International Conference on Quantitative Genetics*, edited by B. S. WEIR, M. M. GOODMAN, E. J. EISEN and G. NAMKOONG. Sinauer Associates, Sunderland, MA.

SPICKETT, S. G., and J. M. THODAY, 1966 Regular responses to selection 3. Interaction between located polygenes. Genet. Res. **7:** 96–121.

STUBER, C. W., and M. EDWARDS, 1986 Genotypic selection for improvement of quantitative traits in corn using molecular marker loci. Proc. 41st Annual Seed Trade Assoc. **41:** 70–83.

TANKSLEY, S. D., and J. D. HEWITT, 1988 Use of molecular markers in breeding for soluble solids in tomato—a re-examination. Theor. Appl. Genet. **75:** 811–823.

TANKSLEY, S. D., H. MEDINA-FILHO and C. M. RICK 1981 The effect of isozyme selection on metric characters in an interspecific backcross of tomato—basis of an early screening procedure. Theor. Appl. Genet. **60:** 291–296.

TANKSLEY, S. D., J. C. MILLER, A. H. PATERSON, and R. BERNATZKY, 1988 Molecular mapping of plant chromosomes, pp. 157–173 in *Proceedings of the 18th Statler Genetics Symposoium*, edited by J. P. GUSTAFSON, Plenum Press, New York.

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.

TAZIMA, Y., 1964 *The Genetics of the Silkworm*. Prentice-Hall, Englewood Cliffs, N.J.

THODAY, J. M., 1961 Location of polygenes. Nature **191:** 368–370.

WRIGHT, S., 1968 *Evolution and the Genetics of Populations*. University of Chicago Press, Chicago.

YOUNG, N. D., and S. D. TANKSLEY, 1988 Restriction fragment length polymorphism maps and the concept of graphical genotypes. Theor. Appl. Genet. **77:** 95–101.

Communicating editor: B. BURR