Euphytica **125:** 69–79, 2002.

The Role of epistasis in controlling seed yield and other agronomic traits in an Andean × Mesoamerican cross of common bean (*Phaseolus vulgaris* L.)

William C. Johnson[1] & Paul Gepts*

Department of Agronomy and Range Science, University of California, One Shields Avenue, Davis, CA 95616-8515, U.S.A.; [1]*Present address: Seminis Vegetable Seeds, 37437 Highway 16, Woodland, CA 95695, U.S.A.;* **Author for correspondence; e-mail: plgepts@ucdavis.edu*

Received 2 August 2000; accepted 5 August 2001

Key words: digenic interactions, linkage map, inbreeding depression, *Phaseolus vulgaris*, QTL

Summary

Epistasis is a pervasive phenomenon in biology. Nevertheless, attempts at identifying epistatic interactions with quantitative trait loci (QTL) analyses have yielded inconsistent results. In this study, we attempt to determine the genetic control of outbreeding depression and the possible role of epistasis following a wide cross in common bean (*Phaseolus vulgaris* L.). A recombinant inbred population, derived from a cross between Andean and Mesoamerican common bean cultivars, was evaluated in two markedly contrasting environments. A low-density linkage map based on AFLPs was used to locate QTLs for the number of days to maturity, average daily biomass and seed yield accumulation, and harvest index. Both independently acting and digenic epistatic QTLs of similar magnitude were identified. A majority of the loci involved in these epistatic interactions did not have an independent effect. Although we did find evidence for strong epistatic control of the traits investigated, we also found, in contrast to other recent studies, that there was no evidence for a bias toward coadapted gene complexes at the level of digenic epistasis. We discuss these results in relation to the role of epistasis in the evolutionary history of the species and methodological difficulties in detecting epistasis.

Abbreviations: AFLP – amplified length fragment polymorphism; BIO - biomass; BYD – biomass yield per day; CY – California Dark Red Kidney × Yolano; DTM – number of days to maturity; EQTL – epistatically-acting quantitative trait loci; HI – harvest index; IQTL – independently-acting quantitative trait loci; QTL – quantitative trait locus; RFLP – restriction fragment length polymorphism; RI – recombinant inbred; SCAR – sequence characterized amplified region; SYD – seed yield per day; YIE – yield

Introduction

One major difference in the contributions of Fisher and Wright to the neo-Darwinian theory of evolution is the role attributed to epistasis in the differentiation of natural populations. Fisher believed that selection during evolution acted primarily on the effect of individual genes, independent of their interactions with other genes. Wright, on the other hand, proposed a major role for epistatic gene action in the development an adaptive landscape with fitness peaks consisting of co-adapted gene complexes (Whitlock et al., 1995; Fenster et al., 1997). Epistatic interactions among genes can take on many forms (Frankel & Schork, 1996; Fenster et al., 1997). For example, evolutionary geneticists posit that the reproductive isolation between species arise through the accumulation of complementary genes that have no effect within a taxon, but which have a deleterious phenotypic effect when combined with genes from other taxa (Lynch, 1991; Orr, 1995; Hutter, 1997).

Outbreeding depression, i.e., the reduced average fitness of progeny from wide crosses, has been attributed to a breakup of coadapted gene complexes or preferred epistatic relationships (Mayr, 1963; Templeton, 1981). Yet, the results of QTL studies have been inconsistent as to the importance of epistasis (Tanksley, 1993). A recent study (Li et al., 1997) on inter-gene pool derived populations derived from crosses between *indica* and *japonica* cultivars of cultivated rice (*Oryza sativa* L.) using markers throughout the genome found that the performance of grain yield components was conditioned by high levels of digenic epistatic interactions in addition to main effects of individual loci. The effects of recombinant genotypes were predominantly negative suggesting that different co-adapted gene complexes controlling yield had arisen in the *indica* and *japonica* gene pools.

Common bean (*Phaseolus vulgaris* L.) presents an interesting model to study the inheritance of outbreeding depression in wide crosses and the possible role of epistasis in speciation. This species consists of two diverged geographical gene pools (Mesoamerican and Andean) (Singh et al., 1991; Gepts, 1993, 1998). Progeny from crosses between these gene pools may suffer from a number of phenotypic abnormalities, beginning with F_1 hybrid weakness in some genotypes (Shii et al., 1980; Gepts & Bliss, 1985; Koinange & Gepts, 1992). In the F_2 and later generations additional abnormal segregants appear, including crippled seedlings exhibiting virus-like symptoms and/or variegated leaves (Singh & Molina, 1996), partial to complete male sterility (Sprecher & Khairallah, 1989), and diminished seed yield potential (Patiño & Singh, 1989; Singh et al., 1989; Johnson & Gepts, 1999). The simple genetic control of some of these abnormalities (Shii et al., 1980; Singh & Molina, 1996) suggests that *P. vulgaris* may be in the process of incipient speciation into an Andean and a Mesoamerican species.

In this paper, we present results of a QTL analysis in a recombinant inbred (RI) population resulting from an Andean × Mesoamerican cross in common bean. The performance of this population was on average below that of the lower parent. In addition, the performance of the best RI lines in this population did not exceed that of the best parent (Johnson, 1997; Johnson & Gepts, 1999). Thus, the behavior of this population is typical of what has been observed for the progenies of Andean × Mesoamerican crosses (e.g., Kornegay et al., 1992; Welsh et al., 1995). This pattern of quantitative trait inheritance suggested a possible role for epistatic interactions. Two working hypotheses were considered. The 'co-adaptation' hypothesis proposes that superior performance is due to the presence – in each of the gene pools – of a unique suite of genes, the expression of which is carefully regulated both in time and space. The alternative hypothesis, which we label the 'hopeful recombinant' hypothesis, also proposes that performance depends on the carefully regulated expression of a suite of genes. This suite would be similar between the two gene pools rather than be unique to either of them. However, because the number of genes in such a suite would likely be quite large for a trait as complex as yield, the likelihood of recovering this suite in the progeny of a wide cross would be quite low.

The two hypotheses can be distinguished by the nature of the parental alleles at performance loci in high- vs. low-yielding progenies. In support of the co-adaptation hypothesis, one would observe a majority of alleles of either the maternal or paternal parent at epistatically interacting loci conditioning performance-related traits in high-yielding lines, i.e. one would observe primarily parental types among high-yielding lines. Under the hopeful recombinant hypothesis, one would expect alleles of both parents at performance loci in each high-yielding line, i.e. a high proportion of recombinant types of epistatic interactions conditioning superior performance at the level of digenic epistasis. In contrast with the results of Li et al. (1997) in rice, we found that co-adapted gene complexes apparently did not play a predominant role in determining performance of our population derived from an inter-gene pool cross in common bean, although digenic epistatic interactions clearly played an important role.

Materials and methods

Plant material and field trials

A RI population – California Dark Red Kidney (Andean) × Yolano (Mesoamerican) (CY) – was used to study the relationship between patterns of inheritance and performance of inter-gene pool hybrid populations (Johnson & Gepts, 1999). The parents of this population are representative of their respective gene pools based on molecular marker and phenotypic data (Singh et al., 1991), and are commonly used cultivars in California. The development of the CY population, consisting of 150 F_7 recombinant inbred lines (RILs), has been described before (Johnson & Gepts, 1999).

The CY population was evaluated in trials at Davis and Salinas in 1995. Salinas represents a cool coastal climate near the Monterey Bay with average July temperature of 19 °C. Davis, in the Sacramento Valley, has a warm summer climate with average July temperature of 26 °C. The RI population was replicated three times in each trial. Ten seeds for each RIL were space planted per 1 m plot, with a two plant in-row border between plots. Standard agronomic practices were maintained at each site. Days to maturity (DTM) were measured as the number of days from planting (day 0) to the first day when half of the pods on half of the plants in the plot were dry. Seed yield per day (SYD) was calculated as seed yield (YIE) divided by DTM. Total aboveground dry weight (BIO) was measured by harvesting the entire plot (sheared at ground level, without attempt to collect fallen leaves) 5 to 10 days after maturity, thoroughly drying the plants in a drying shed (10 to 20 days), and weighing. Biomass yield per day (BYD) was calculated as BIO divided by DTM, and Harvest Index (HI) was calculated as YIE divided by BIO.

Genetic markers

DNA was isolated from the retained leaf samples as described in Gepts et al. (1992). DNA was quantified using a Hoefer TKO 100 DNA Fluorimeter (Hoefer Scientific Instruments, San Francisco, CA). RAPD reactions were performed in an Ericomp Twinblock thermal cycler (Ericomp, San Diego, CA). Reaction parameters were similar to those of Williams et al. (1990), but with 10–20 ng of genomic DNA in a total reaction volume of 25 μl. The thermal profile was 1 cycle of 2 min at 94 °C; 3 cycles of 1 min at 94 °C / 1 min at 35 °C / 2 min at 72 °C; 32 cycles of 10 s at 94 °C / 30 s at 35 °C / 1 min at 72 °C; and 1 cycle of 5 min at 72 °C. Seed proteins were extracted as in Gepts et al. (1986) and one-dimensional SDS / PAGE electrophoresis was performed as in Ma & Bliss (1978). Isozymes were analyzed and scored using the methods of Koenig & Gepts (1989). RFLP analysis was performed using previously mapped probes developed by Nodari et al. (1992) and Vallejos et al. (1992). Probes were radiolabeled using α-[^{32}P]dCTP (Amersham) by the random priming method (Feinberg & Vogelstein, 1984). Southern hybridization was performed according to the Zetabind protocol (AMF-CUNO, Meriden CT). Probes were hybridized to the membranes in a solution of 4X SSPE and 7.5% SDS at 65 °C overnight, followed by two washes with 2X SSC 0.1% SDS at 60 °C for 15 minutes and one or two washes in 0.1X SSC 0.1% SDS at 60 °C for 30 minutes. X-ray film was exposed to membranes for 1 to 15 days. SCARs (sequence characterized amplified regions) used as molecular markers for mapping were analyzed with the conditions suggested by their developers (Adam-Blondon et al., 1994; Johnson et al., 1997). AFLP analysis was performed using the protocol of Vos et al. (1995), with modifications described in Johnson (1997). The *fin* gene, which conditions determinacy vs. indeterminacy in common bean (Leakey 1988), was scored at flowering.

Mapping

Identification of linkage groups was performed using Mapmaker 3.0b (Lander et al., 1987) according to Menéndez et al. (1997). Heterozygotes, where observable (RFLPs, isozymes, some SCARs and AFLPs, and morphological markers), were noted but were considered missing data for the purposes of mapping and QTL analysis. Linkage groups were anchored using previously mapped RFLP, morphological, seed protein, isozyme, or SCAR markers (Freyre et al., 1998). From the 196 markers mapped, 79 framework markers were chosen based on marker information content and map location to maximize genome coverage at an optimal spacing of 10–15 cM. Marker information (marker names, fragment sizes, linkage orientations, segregation distortion, and flanking markers) is available in Johnson (1997).

Single-factor QTL analysis

Standard interval mapping and simplified composite interval mapping for IQTL (independently acting or potentially monogenic QTL) identification was performed using MQTL (Tinker & Mather, 1995a). IQTL were declared significant based on simplified interval mapping (Tinker & Mather, 1995b) with a 10% experimentwise error rate estimated by 1000 permutations (chosen for computational feasibility) of the data sets (Doerge & Rebai, 1996). Linked IQTL were defined as independent where simplified composite interval mapping peaks previously identified as statistically significant with standard interval mapping were observed at a distance larger than 20 cM (Tinker & Mather, 1995b). Approximate R^2 values were calculated from test statistics as described by Tinker (1996). Multiple regression using all significant IQTL as the model was performed using MQTL (Tinker & Mather, 1995b).

Epistatic QTL analysis

SAS PROC GLM (SAS, 1988) was used to survey the regions likely to be involved in two-locus epistasis for the traits of interest. For purposes of computational feasibility, initial analyses using widely spaced markers ($\bar{x}$ = 26 cM, n=41) spaced throughout the genome were used to identify regions with interaction effects ($p < 0.001$) on DTM, SYD, BYD, and HI. Additional SAS PROC GLM marker interaction analyses with increased marker density in the regions identified as significant in the first round were used to more clearly delineate the genomic regions with significant epistatic effects. For cases where two or more significant interactions appeared to be due to linkage (by examining the F statistic profiles along the two genomic regions) only the pairs with the strongest associations (lowest P>F) were maintained for further analysis.

Very highly significant epistatic interactions ($p < 0.001$) identified with the genomic scans using SAS were chosen as putative EQTL (digenic epistatic QTL) to initiate genome scans using MQTL. Linked EQTL were defined as independent where interval mapping peaks were observed at a distance of more than 25 cM. Successive rounds of MQTL interaction scans were used to localize interactions to 5 cM intervals by maximizing the interaction model (a + b + a*b) vs. additive effects model (a + b) test statistic (N. Tinker, pers. comm.) for both the anchor marker (a) and the interacting marker (b). EQTL identified as significant for a trait*environment at an approximate 10% experimentwise error rate – determined through 1000 permutations (chosen for computational feasibility) of anchor marker*trait*environment specific data sets (Doerge & Rebai, 1996) – were analyzed for significance for the same trait in other environments and for the other traits in the same environment. Successive rounds of two-way interaction scans were performed until no further interactions were identified. Reported results include only EQTL significant at a maximum 5% experimentwise error rate. For comparison of significance, among the 6 CY RI population EQTL where MQTL and SAS analyses are nearly equivalent (cases where MQTL estimates both interacting loci to be 0 cM from the nearest marker and there are <40 genotypes with missing marker data), the EQTL exhibiting a test statistic greater than the 5% MQTL experimentwise error rate threshold all have a mean probability according to the SAS Proc GLM of 0.0001 (W.C. Johnson, unpublished results). Approximate EQTL R^2 values are calculated from test statistics as described by Tinker (1996). Analysis of variance and mean separations of epistatic marker classes (2 parental, 2 recombinant) were performed using SAS Proc Means (SAS, 1988). Higher level interactions (3 locus and 4 locus models) suggested by the results of two-way interactions were analyzed using a Duncan's multiple range test and Proc GLM of SAS (SAS, 1988).

Results

Linkage map

The linkage map consists of 11 large linkage groups, numbered from B(ean)1 to B11 in accordance with the common bean framework map (Freyre et al., 1998) (Figure 1). Three additional, smaller linkage groups, designated with letters instead of numbers, could not be related to previously described linkage groups. Genome coverage was 862 cM (Kosambi), compared to previous reports in common bean of 960 (Vallejos et al., 1992) and 827 (Nodari et al., 1993). Effective genome coverage (within 15 cM of a framework marker) for QTL analysis in this population is at minimum 967 cM, because the present map includes anchor markers known to map to the distal portions of previously described linkage groups, and because the unassigned linkage groups also likely represent portions of the 11 assigned linkage groups (2n=2x=22). Of the 196 markers scored, 192 were mapped to a linkage group. AFLP markers tended to map to a cluster (multiple markers within 5 cM) on each linkage group, usually near the center, presumably representing centromeric regions in which recombination is reduced relative to the physical length of the chromosome. On average, 10.1 markers were observed per cluster in nine linkage groups (not observed in B7 or B11). The markers shown in Figure 1 represent the subset of the markers chosen to maximize genome coverage and minimize missing data that were used for QTL analysis (the framework markers). Roughly one third of the markers showed significant segregation distortion based on chi-square tests, with approximately equal numbers of marker loci skewed toward the Mesoamerican or the Andean parental genotypes (Figure 1). Three AFLP markers displayed complete or near-complete segregation distortion toward the Yolano allele, and consequently could not be mapped.

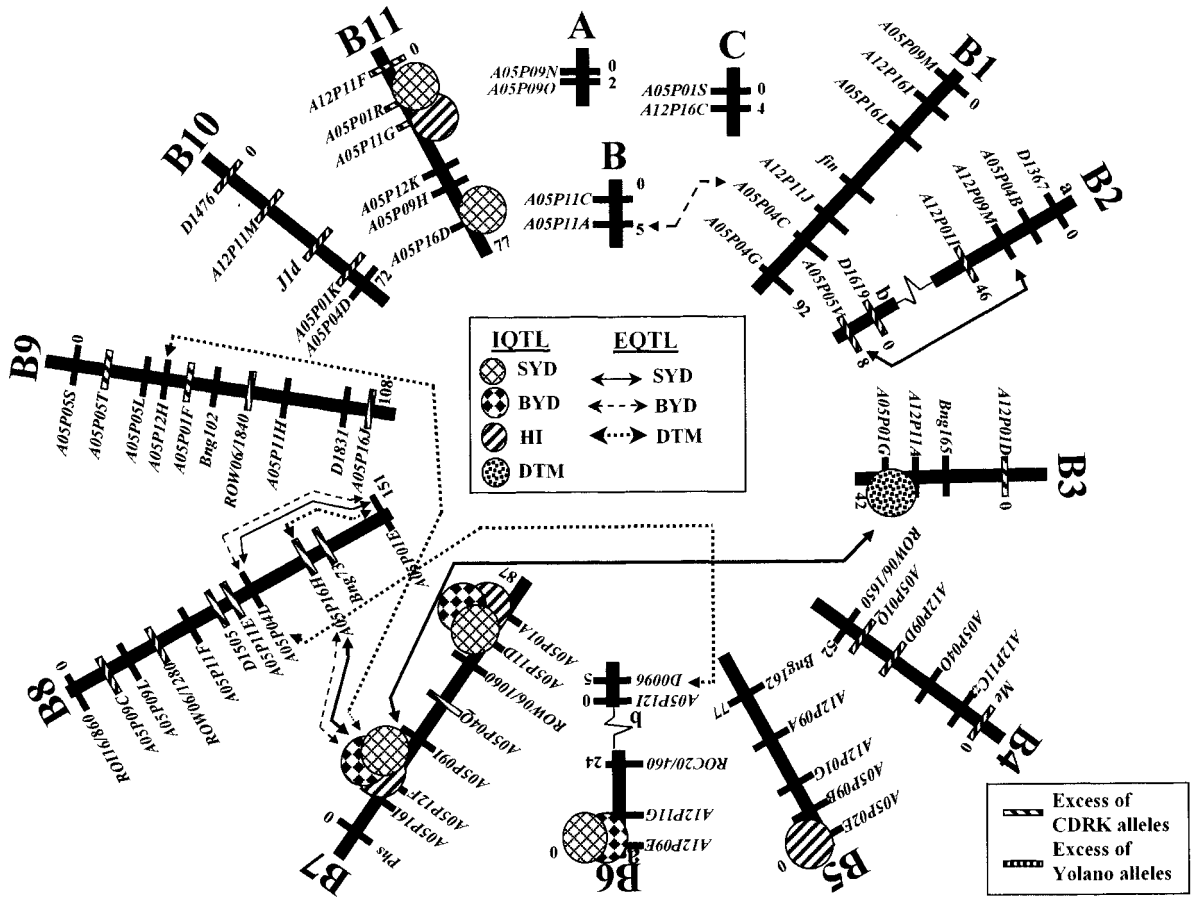

Figure 1. Linkage map and QTL location for across-location averages of DTM (number of days to maturity), SYD (seed yield per day), BYD (biomass yield per day), and HI (harvest index) in the CY RI population. Thick black bars represent linkage groups. Framework markers, the subset of markers chosen to maximize genome coverage with the highest information content, are represented by cross bars and locus name. Numbers at the right of each linkage group represent genetic distances in Kosambi cM. Crossbars with pattern fills represent framework markers exhibiting segregation distortion. IQTL (independently acting QTL) are represented by patterned circles located at the locus with the highest likelihood, and EQTL (epistatically interacting QTL) are represented by patterned lines with arrowheads pointing to the loci with the highest likelihood. Marker information (marker types, marker names, linkage phases, and segregation ratios) is available in Johnson (1997).

Independent gene action

Agronomic performance of the RI population was inferior to the parents for all traits measured and in all field trials. None of the recombinant inbred lines was significantly superior to the best parent for any of the traits measured (Johnson, 1997; Johnson & Gepts, 1999). Climatic conditions between the locations are more different than year-to-year variation within location. Thus, a comparison of the performance at Davis and Salinas is a more stringent test of the effect of environmental variation than multi-year test within any of these locations. Significant effects, as determined by MQTL (Tinker & Mather, 1995a), were identified for IQTL (independently acting QTLs), when individual locations, years, and general averages were considered for the four traits of interest: days to maturity (DTM), seed yield per day (SYD), biomass yield per day (BYD), and harvest index (HI). The average QTL effect was 8% for the four traits of interest. No IQTL were detected for DTM or HI at Salinas (Table 1).

Epistatic gene action

Twenty-two epistatic interactions were detected (summarized in Table 2 and Figure 1) for the four traits of interest. Each of the interactions accounted on average

Table 2. Summary of epistatically-acting quantitative trait loci detected in the California Dark Red Kidney × Yolano recombinant inbred population of common bean

Trait[a]	Locus pair		R^2 (%)	Phenotypic average[b]				Phenotypic average		Highest performing marker class[c]	
	A	B		CC	CY	YC	YY	Parental class[b]	Recombinant class[b]	P vs. R	Allelic classes
Davis 1995											
DTM	B1 – *A05P09M*[d]	B1 – *A05P04C*	9*,[e]	101	104	104	97	101	103	ns	CC, CY, YC
	B7 – *A05P12F*	B9 – *A05P12H*	13**	103	105	104	98	100	104	P	YY
BYD	B1 – *fin*	B2 – *D1619*	9*	3.9	3.3	2.8	3.7	3.1	3.2	ns	CY, YC, YY
	B1 – *A05P04C*	**B – *A05P11C***	9*	3.4	3.6	2.4	4.1	3.6	3.2	ns	CY, YY
Salinas 1995											
SYD	B2 – *A12P09M*	B2 – *D1619*	10*	2.1	2.7	2.8	2.3	2.2	2.6	R	CY, YC, YY
	B3 – *A05P01G*	**B7 – *A05P09I***	10*	2.5	2.5	2.0	2.8	2.7	2.3	P	YY
	B6 – *A12P09E*	B9 – *A05P11H*	9*	2.9	2.4	2.0	2.5	2.8	2.4	P	CC, CY, YY
	B7 – *A05P12F*	B8 – *A05P16H*	14*	2.0	2.7	2.7	2.3	2.1	2.6	R	CY, YC, YY
BYD	B1 – *fin*	B8 – *A05P11F*	10*	4.7	5.5	5.3	4.5	5.2	4.9	ns	ns
	B3 – *A05P01G*	**B7 – *A05P09I***	12**	5.2	5.1	4.3	5.5	5.3	4.7	P	CC, CY, YY
	B7 – *A05P12F*	B8 – *A05P16H*	12**	4.6	5.2	5.4	4.6	4.4	5.2	R	CY, YC
HI	B1 – *fin*	B6 – *A12P09E*	13**	0.5	0.5	0.5	0.4	0.5	0.5	ns	CC, CY, YC
	B1 – *A05P01R*	B3 – *A12P01A*	10*	0.5	0.5	0.5	0.4	0.5	0.5	R	CC, CY, YC
Average 1995											
DTM	B6 – *D0096*	B8 – *A05P04L*	9*	107	101	104	106	104	104	ns	ns
	B7 – *A05P12F*	B9 – *A05P12H*	13**	105	107	106	101	103	107	P	YY
SYD	B2 – *A12P09M*	B2 – *A05P05V*	10*	1.2	1.7	1.7	1.5	1.3	1.7	R	CY, YC, YY
	B3 – *A05P01G*	**B7 – *A05P09I***	10*	1.6	1.6	1.2	1.9	1.7	1.4	P	CY, YY
	B7 – *A05P12F*	B8 – *A05P16H*	10*	1.2	1.7	1.8	1.5	1.3	1.7	R	CY, YC, YY
	B8 – *A05P04L*	B8 – *A05P01E*	9*	1.4	1.6	2.1	1.4	1.4	1.7	R	CY, YC
BYD	B1 – *A05P04C*	B – *A05P11A*	9*	4.4	4.3	3.3	4.6	4.4	4.0	ns	CC, CY, YY
	B7 – *A05P12F*	B8 – *A05P16H*	9*	3.8	4.3	4.6	3.9	3.7	4.3	R	CY, YC, YY
	B8 – *A05P04L*	B8 – *A05P01E*	11**	4.0	4.6	5.0	3.7	4.0	4.5	R	CY, YC

[a] BYD: biomass yield per day; DTM: days to maturity; SYD: seed yield per day; HI: harvest index.
Marker classes are designated with one letter representing each locus because RILs are predominantly homozygous. Means calculated by MQTL.

[b] Marker classes are parental (CC + YY) or recombinant (CY+YC). Means calculated by SAS using nearest flanking markers and omitting genotypes with missing marker data.

[c] Marker classes (parental *versus* recombinant in the first column, four individual allelic classes in the second) exhibiting statistically significant (0.05 level) higher trait expression according to Duncan's multiple range test using nearest flanking markers and analyzed with SAS. MQTL does not perform multiple range tests and SAS does not infer missing marker data or perform interval analysis. Therefore the subset of marker classes showing statistically significant differences according to SAS is slightly different from those identified by MQTL. Ns: non-significant differences.

[d] Linkage group – framework marker (see Figure 1). MQTL markers displayed in bold also exhibit a significant test statistic (0.10 experimentwise error rate) with simple interval mapping for IQTL.

[e] **, *: significant at the $p = 0.01$ and $p = 0.05$ levels, respectively.

for 10% of the variation in the traits. Eight interactions included a locus that also had a significant effect as an IQTL. Nine of the interactions involved EQTLs linked to framework markers *A05P12F* and *A05P09I*, which were also identified as having the most highly significant effect as IQTLs for several traits. Two of these interactions had an effect on DTM although none of the component loci of these two epistatic interactions were identified as an IQTL.

Discussion

Role of epistasis

Our results show that epistatic gene interactions play an important role – in addition to independent gene action – in determining performance in the wide cross analyzed. The two types of gene action were similar with respect to the number of interactions or loci

Table 3. Comparison of the performance of parent-like and recombinant lines in the California Dark Red Kidney × Yolano recombinant inbred population of common bean

Trait	CDRK skewed: P>F	Yolano skewed: P>F	Either skewed: P>F	Parental RILs mean	Recombinant RILs mean
Davis 1995					
DTM	0.90	0.01*	0.07	100	102
SYD	0.94	0.05	0.21	0.77	0.59
BYD	0.84	0.09	0.36	3.63	3.30
HI	0.90	0.26	0.40	0.18	0.16
Salinas 1995					
DTM	0.43	0.06	0.05*	106	108
SYD	0.19	0.22	0.07	2.70	2.39
BYD	0.10	0.55	0.07	5.38	4.91
HI	0.64	0.23	0.34	0.50	0.48
Average 1995					
DTM	0.72	0.01*	0.03*	103	105
SYD	0.44	0.08	0.08	1.74	1.49
BYD	0.48	0.19	0.15	4.50	4.11
HI	0.33	0.17	0.28	0.34	0.32

Table 4. Summary of independently- and epistatically-acting quantitative trait loci identified in the California Dark Red Kidney × Yolano recombinant inbred population of common bean

	Trait				
	DTM	SYD	BYD	HI	All
IQTLs					
Number of IQTLs identified	1	5	3	4	13
Mean R^2 (%) explained by individual IQTL	8	7	7	7	7
R^2 (%) collectively explained by all IQTL	7	27	19	18	18
Digenic epistatic interactions (EQTLs)					
Number of interactions detected	2	4	3	0	9
Mean R^2 (%) explained by individual EQTL pairs	11	10	10	NA [a]	10
No. interactions conditioned by 2 IQTLs (any trait)	0	1	0	NA	1
No. interactions conditioned by a IQTL (any trait) interacting with a non-IQTL	1	1	1	NA	3
No. interactions conditioned by non-IQTL (any trait)	1	2	2	NA	5
Minimum no. loci involved in interactions	4	6	5	NA	9
Mean multiepistativity of individual loci (Li et al., 1997)	1	1.3	1.2	NA	2.0

[a] NA: not applicable.

involved and the mean R^2 per independent locus or epistatic interaction (Table 4). Although the number of epistatic interactions tested was much larger than the number of independent actions [n(n-1)/2 vs. n, where n is the number of marker loci], the choice of a similar genome- or experimentwise error rate (5%) insured that the numbers of significant tests were comparable between the independent and epistatic action analyses. Whereas experimental conditions, in particular population size, only allowed us to investigate digenic epistasis, a 'chain-like' relation among digenic interactions suggested that higher-order interactions may also be important. For example, the two-way interactions involving loci *A05P01G* (B3),

Table 1. Summary of independently acting quantitative trait loci detected in the California Dark Red Kidney × Yolano recombinant inbred population of common bean

Trait[a]	Map location		Gene action	
	Linkage group	Nearest framework marker[b]	R^2 (%)	Additive effect[c]
Davis 1995				
DTM	3	*A05P01G*	14	4.7
SYD	5	*A05P09B*	10	0.35
	6	*A12P09E*	10	0.38
	7	*A05P12F*	7	–0.29
	7	*A05P11D*	12	–0.41
	11	*A05P01R*	6	0.28
	11	*A05P16D*	2	0.18
BYD	6	*A12P09E*	11	0.83
	7	*A05P12F*	3	–0.42
	7	*A05P11D*	16	–1.01
	B	*A05P11C*	10	–0.76
HI	5	*A05P02E*	13	0.09
	7	*A05P12F*	7	–0.06
	7	*A05P11D*	11	–0.09
	11	*A05P11G*	9	0.07
Salinas 1995				
SYD	7	*A05P12F*	4	–0.27
	7	*A05P11D*	6	–0.34
	11	*A05P09H*	11	0.49
BYD	7	*A05P12F*	2	–0.29
	7	*A05P11D*	7	–0.55
	9	*A05P12H*	11	0.72
Average 1995				
DTM	3	*A05P01G*	8	3.3
SYD	6	*A12P09E*	8	0.37
	7	*A05P12F*	5	–0.28
	7	*A05P11D*	9	–0.37
	11	*A05P01R*	5	0.29
	11	*A05P16D*	6	0.30
BYD	6	*A12P09E*	8	0.65
	7	*A05P12F*	3	–0.35
	7	*A05P11D*	11	–0.78
HI	5	*A05P02E*	7	0.06
	7	*A05P12F*	5	–0.04
	7	*A05P11D*	6	–0.05
	11	*A05P11G*	8	0.06

[a] DTM: Number of days to maturity; SYD: Seed yield per day; BYD: biomass yield per day; HI: harvest index.
[b] See Figure 1.
[c] Phenotypic effect of putative IQTL locus: measured in units of the trait, where positive and negative values indicate the female-derived and male-derived alleles, respectively, contributing factors for higher trait expression.

A05P12F-A05P09I (B7), *A05P04L-Bng73* (B8), and *A05P01E* (B8) (Figure 1) could actually be involved in a four-way interaction. Additional evidence for these higher-order interactions could be obtained by examining larger populations. Several recent papers have also reported on important epistatic effects detected by QTL analysis (Long et al., 1995; Holland et al., 1997; Li et al., 1997).

The important role of non-IQTL loci in epistatic interactions (Table 4) has a bearing on genetic research areas, such as marker-assisted selection. Marker-assisted selection has been performed so far for IQTL loci, i.e., genes that have a measurable independent effect, be they major genes or QTLs. Our results provide opportunities for testing of marker-assisted selection in Andean × Mesoamerican crosses. Genomic regions carrying IQTLs or EQTLs could be transferred into the genomic background of the appropriate parent. The transfer of a limited number of genomic regions, individually or in combination, by marker-assisted selection could alleviate the problem associated with low probability of recovery of superior recombinants.

Speciation and epistasis in common bean

P. vulgaris is a species that may be undergoing allopatric speciation. Evidence for this speciation process is based primarily on the existence of partial F_1 hybrid weakness and the degree of molecular divergence between the Andean and Mesoamerican gene pools (Gepts & Bliss, 1985; Koenig & Gepts, 1989; Koinange & Gepts, 1992). To account for the reduced productivity of the progeny of Andean × Mesoamerican crosses, we had proposed two, not mutually exclusive hypotheses, namely the co-adaptation and hopeful recombinant hypotheses. The co-adaptation hypothesis implies that progeny with superior viability, fertility, and performance, will show parental combinations of alleles in their epistatic interactions. Individuals with recombinant combinations of alleles in epistatic interactions are expected to have inferior performance.

The importance of parental vs. recombinant epistatic interactions in promoting higher performance can be evaluated in various ways. In general, hybrid lethality is thought to be due to epistatic interactions among genes from the parental gene pools when these are united in a hybrid, whereas on their own (i.e., in their respective gene pools) they do not affect viability (Dobzhansky, 1937; Muller, 1940). The population studied here did not show the F_1 hybrid inviabil-

ity caused by the complementary *Dl* genes (Gepts & Bliss, 1985). However, in subsequent generations, there was circumstantial evidence for the presence of (recessive) lethality factors. Some 18% of the F_2 generation died soon after germination. Three AFLP markers showed complete or near-complete distortion with one of the parental alleles almost completely or entirely fixed in the population. In addition, a higher than expected proportion of RI lines were predominantly parental genotypes. Predominantly parental genotypes were those showing a statistically significant excess of either maternal or paternal alleles based on a chi-square test using all framework markers and a $p = 0.05$ level of significance. Conversely, recombinant lines contained a similar proportion of maternal and paternal alleles, i.e. no significant difference between the number of maternal and paternal derived markers. In the RI population, 18 and 16 RI lines consisted primarily of CDRK and Yolano alleles, respectively. At the $p = 0.01$ level, 6 and 5 RI lines showed predominantly CDRK and Yolano alleles, respectively, which is over seven times the frequency expected by chance.

Further evidence for allelic co-adaptation comes from a comparison of the performance of the subset of the RI lines with predominantly parental genotypes to that of the predominantly recombinant genotypes. As Table 3 shows, the subset of RI lines with predominantly parental genomes outperforms the remaining RI lines for all traits. Caution must be exercised in interpreting these results, as only 18 CDRK-like and 16 Yolano-like RI lines were used in this analysis.

In contrast, 9 of 15 interactions associated with higher performance were recombinant allele combinations in the entire RI population (Table 2). These data do not necessarily contradict the evidence mentioned in previous paragraphs, which ascribe an important role for parental epistatic interactions. An analysis of recombinant RI lines shows that some have a performance similar to the best parent-like lines (data not shown). This observation suggests, in turn, that RI lines may carry specific recombinant epistatic interactions that lead to higher performance.

Taken as a whole, our data then do not show a marked advantage for parental epistatic combinations. This observation in turn suggests that the reduced performance observed in Andean × Mesoamerican crosses is not exclusively due to the break-up of co-evolved gene complexes. The existence of superior recombinant allele combinations in addition to superior parental allelic combinations and individual alleles originating in both parents suggests that this reduced performance may be due to the low probability of recovering genotypes with all the necessary genes to achieve high performance given the population size and mating system used in this experiment (see below).

A similar study performed by Li et al. (1997) on the genetics of yield components in rice (*Oryza sativa* L.) arrived at different conclusions regarding the importance of parental epistatic interactions. The comparison with rice is of particular interest because, like common bean, this species consists of two diverged gene pools (*japonica* and *indica*). In contrast with common bean, Li et al. (1997) observed that parental interactions were more likely to increase trait expression in yield components than recombinant interactions. One can speculate that this difference between the two species might be due to a longer duration of divergence between the two major gene pools in rice compared to common bean. The longer divergence time would have provided more opportunities for the development of co-adapted gene complexes (Whitlock et al., 1995). Our observations also substantiate Orr's (1995) speciation model, which states that interspecific hybridization could be hampered by any number of genes and result in any level of decrease fitness. Thus, species differences can become established that involve both parental and recombinant allele epistatic interactions.

Acknowledgements

This research was funded by the US AID Bean/Cowpea CRSP Grant DAN-1310-G-88-6008-00. Special thanks to Robert Lewellen, Jack McBride, and Sharon Bentzen of the USDA-ARS station at Salinas CA for providing field space and assistance. Thanks to Luís Sanchez, Donald Helms, Pedro Pereira, and Steve Temple for agronomic advice and assistance. We would also like to thank Richard Barrick, Andrew Brosnan, Tara Campo, Christhiam Cano, Sylvaine Coulibaly, Mattheus Dahlberg, Douglas Denotter, Robin Emig, Denise Flanahan, Rosanna Freyre, Leslie Goldberg, Carrie Hamlin, Christa Haney, Yuni Jang, Benjamin Johnson, Christopher Johnson, Katharene Johnson, Janos Kovacs, Cristina Menéndez, James Osher, George Pavana, Bernadette Paul, Pedro Pereira, David Posner, Asgar Shirmohamadali, Hau Truong, Phuong Truong, and Patrick Tse, for assistance with fieldwork.

References

Adam-Blondon, A., M. Sévignac, H. Bannerot & M. Dron, 1994. SCAR, RAPD and RFLP markers tightly linked to a dominant gene (*Are*) conferring resistance to anthracnose in common bean. Theor Appl Genet 88: 865–870.

Avery, L. & S. Wasserman, 1993. Ordering gene function: the interpretation of epistasis in regulatory hierarchies. Trends Genet 8: 312–316.

Cheverud, J. & E. Routman, 1996. Epistasis as a source of increased additive genetic variance at population bottlenecks. Evolution 50: 1042–1051.

Dobzhansky, T., 1937. Genetics and the Origin of Species. Columbia Univ. Press, New York.

Doerge, R. & A. Rebai, 1996. Significance thresholds for QTL interval mapping tests. Heredity 76: 459–464.

Edwards, M., T. Helentjaris, S. Wright & C. Stuber, 1992. Molecular-marker facilitated investigations of quantitative trait loci in maize. 4. Analysis based on genome saturation with isozyme and restriction fragment length polymorphism markers. Theor Appl Genet 83: 765–774.

Eshed, Y. & D. Zamir, 1995. An introgression line population of *Lycoperison pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics 141: 1147–1162.

Eshed, Y. & D. Zamir, 1996. Less-than-additive epistatic interactions of quantitative trait loci in tomato. Genetics 143: 1807–1817.

Feinberg, A.P. & B. Vogelstein, 1984. Addendum: A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Analyt Biochem 137: 266–267.

Fenster, C., L. Galloway & L. Chao, 1997. Epistasis and its consequences for the evolution of natural populations. Trends Ecol Evol 12: 282–286.

Frankel, W. & N. Schork, 1996. Who's afraid of epistasis? Nature Genetics 14: 371–373.

Freyre, R., P. Skroch, V. Geffroy, A.-F. Adam-Blondon, A. Shirmohamadali, W. Johnson, V. Llaca, R. Nodari, P. Pereira, S.-M. Tsai, J. Tohme, M. Dron, J. Nienhuis, C. Vallejos, & P. Gepts, 1998. Towards an integrated linkage map of common bean. 4. Development of a core map and alignment of RFLP maps. Theor Appl Genet 97: 847–856.

Gepts, P., 1993. The use of molecular and biochemical markers in crop evolution studies. Evol Biol 27: 51–94.

Gepts, P., 1998. Origin and evolution of common bean: past events and recent trends. HortScience 33: 1124–1130.

Gepts, P. & F.A. Bliss, 1985. F_1 hybrid weakness in the common bean: differential geographic origin suggests two gene pools in cultivated bean germplasm. J Hered 76: 447–450.

Gepts, P., V. Llaca, R.O. Nodari & L. Panella, 1992. Analysis of seed proteins, isozymes, and RFLPs for genetic and evolutionary studies in *Phaseolus*. In: H.-F. Linskens & J.F. Jackson (Eds.), Modern Methods of Plant Analysis (New Series): Seed Analysis, pp. 63–93. Springer, Berlin.

Gepts, P., T.C. Osborn, K. Rashka & F.A. Bliss, 1986. Phaseolin-protein variability in wild forms and landraces of the common bean (*Phaseolus vulgaris*): evidence for multiple centers of domestication. Econ Bot 40: 451–468.

Greenfield, A. & P. Koopman, 1996. SRY and mammalian sex determination. Curr Topics Dev Biol 34: 1–23.

Hanson, M., B. Gaut, A. Stec, S. Fuerstenberg, M. Goodman, E. Coe & J. Doebley, 1996. Evolution of anthocyanin biosynthesis in maize kernels: the role of regulatory and enzymatic loci. Genetics 143: 1395–1407.

Holland, J., H. Moser, L. O'Donoghue & M. Lee, 1997. QTLs and epistasis associated with vernalization responses in oat. Crop Sci 37: 1306–1316.

Howell, P., D. Marshall & D. Lydiate, 1996. Towards developing intervarietal substitution lines in *Brassica napus* using marker-assisted selection. Genome 39: 348–358.

Hutter, P., 1997. Genetics of hybrid inviability in *Drosophila*. Adv Genet 36: 157–185.

Johnson, W., 1997. Improving the Efficiency of Common Bean (*Phaseolus vulgaris* L.) Breeding Programs Using Molecular Markers. PhD dissertation, University of California, Davis.

Johnson, W. & P. Gepts, 1999. Segregation for performance in recombinant inbred populations resulting from inter-gene pool crosses of common bean (*Phaseolus vulgaris* L). Euphytica 106.

Johnson, W., P. Guzmán, D. Mandala, A. Mkandawire, S. Temple, R. Gilbertson & P. Gepts, 1997. Molecular tagging of the *bc-3* gene for introgression into Andean common bean. Crop Sci 37: 248–254.

Koenig, R. & P. Gepts, 1989. Allozyme diversity in wild *Phaseolus vulgaris*: further evidence for two major centers of diversity. Theor Appl Genet 78: 809–817.

Koinange, E.M.K. & P. Gepts, 1992. Hybrid weakness in wild *Phaseolus vulgaris* L. J Hered 83: 135–139.

Koornneef, M., C. Alonso-Blanco, H. Blankestijn-de Vries, C.J. Hanhart & A.J.M. Peeters, 1998. Genetic interactions among late-flowering mutants of Arabidopsis. Genetics 148: 885–892.

Kornegay, J., J.W. White & O.O. De la Cruz, 1992. Growth habit and gene pool effects on inheritance of yield in common bean. Euphytica 62: 171–180.

Lander, E.S., P. Green, J. Abrahamson, A. Barlow, M. Daly, S.E. Lincoln & L. Newburg, 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1: 174–181.

Leakey, C.L.A., 1988. Genotypic and phenotypic variation in common bean. In: P. Gepts, (Ed.), Genetic Resources of *Phaseolus* Beans, pp. 245–327. Kluwer, Dordrecht, the Netherlands.

Li, Z., S. Pinson, W. Park, A. Paterson & J. Stansel, 1997. Epistasis for three grain yield components in rice (*Oryza sativa* L.). Genetics 145: 453–465.

Li, Z., S. Pinson, J. Stansel & W. Park, 1995. Identification of QTL for heading date and plant height in rice using RFLP markers. Theor Appl Genet 91: 374–381.

Li, Z., S. Pinson, J. Stansel & W. Park, 1995. Characterization of quantitative trait loci contributing to field resistance to sheath blight (*Rhizoctonia solani*) in rice. Theor Appl Genet 91: 382–388.

Long, A., S. Mullaney, L. Reid, J. Fry, C. Langley & T. Mackay, 1995. High resolution mapping of genetic factors affecting abdominal bristle number in *Drosophila melanogaster*. Genetics 139: 1273–1291.

Lynch, M., 1991. The genetic interpretation of inbreeding depression and outbreeding depression. Evolution 45: 622–629.

Ma, Y. & F.A. Bliss, 1978. Seed proteins of common bean. Crop Sci 17: 431–437.

Mayr, E., 1963. Animal Species and Evolution. Columbia Univ. Press, New York.

Menéndez, C., A. Hall & P. Gepts, 1997. A genetic linkage map of cowpea (*Vigna unguiculata*) developed from a cross between two inbred, domesticated lines. Theor Appl Genet 95: 1210–1217.

Merlot, S. & J. Giraudat, 1997. Genetic analysis of abscisic acid signal transduction. Plant Phys 114: 751–757.

Muller, H., 1940. Bearing of the *Drosophila* work on systematics. In: J. Huxley (Ed.), The New Systematics, pp. 185–268. Clarendon, Oxford.

Nodari, R.O., 1992. Towards an integrated linkage map of common bean (*Phaseolus vulgaris* L.). PhD dissertation, University of California, Davis.

Nodari, R.O., S.M. Tsai, R.L. Gilbertson & P. Gepts, 1993. Towards an integrated linkage map of common bean. II. Development of an RFLP-based linkage map. Theor Appl Genet 85: 513–520.

Nodari, R.O., S.M. Tsai, P. Guzmán, R.L. Gilbertson & P. Gepts, 1993. Towards an integrated linkage map of common bean. 3. Mapping genetic factors controlling host-bacteria interactions. Genetics 134: 341–350.

Orr, H., 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics 139: 1805–1813.

Paterson, A.H., S. Damon, J.D. Hewitt, D. Zamir, H.D. Rabinowitch, S.E. Lincoln, E.S. Lander & S.D. Tanksley, 1991. Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. Genetics 127: 181–197.

Patiño, H. & S. Singh, 1989. Visual selection for seed yield in the F_2 and F_3 generations of nine common bean crosses. Annu Rept Bean Improv Coop 32: 79–89.

Ramsay, L., D. Jennings, E. Bohuon, A. Arthur, D. Lydiate, M. Kearsey & D. Marshall, 1996. The construction of a substitution library of recombinant backcross lines in *Brassica oleracea* for the precision mapping of quantitative trait loci. Genome 39: 558–567.

SAS, 1988. SAS/STAT User's Guide, Release 6.03 Edition. SAS Institute, Cary, NC.

Shii, C.T., M.C. Mok, S.R. Temple & D.W.S. Mok, 1980. Expression of developmental abnormalities in hybrids of *Phaseolus vulgaris* L. J Hered 71: 218–222.

Singh, S. & A. Molina, 1996. Inheritance of crippled trifoliolate leaves occurring in interracial crosses of common bean and its relationship with hybrid dwarfism. J Hered 87: 464–469.

Singh, S.P., C. Cajiao, J.A. Gutiérrez, J. García, M.A. Pastor-Corrales & F.J. Morales, 1989. Selection for seed yield in inter-gene pool crosses of common bean. Crop Sci 29: 1126–1131.

Singh, S.P., P. Gepts & D.G. Debouck, 1991. Races of common bean (*Phaseolus vulgaris* L., Fabaceae). Econ Bot 45: 379–396.

Sprecher, S. & M. Khairallah, 1989. Association of male sterility with gene pool recombinants in bean. Annu Rept Bean Improv Coop 32: 56–67.

Tanksley, S., 1993. Mapping polygenes. Annu Rev Genet 27: 205–233.

Templeton, A.R., 1981. Mechansisms of speciation – a population genetic approach. Annu Rev Ecol Syst 12: 23–48.

Tinker, N., 1996. MQTL documentation, version 0.98, ftp://gnome.agrenv.mcgill.ca/pub/genetics/software/MQTL/mqtl.beta098/mqtl.doc.

Tinker, N. & D. Mather, 1995. Methods for QTL with progeny replicated in multiple environments. J Quant Trait Loci 1: http://probe.nalusda.gov:8000/otherdocs/jqtl/jqtl199501/jqtl199515.html.

Tinker, N. & D. Mather, 1995. MQTL: software for simplified composite interval mapping of QTL in multiple environments. J Quant Trait Loci 1: http://probe.nalusda.gov:8000/otherdocs/jqtl/jqtl1995–8002/jqtl8016r8002.html.

Vallejos, E.C., N.S. Sakiyama & C.D. Chase, 1992. A molecular marker-based linkage map of *Phaseolus vulgaris* L. Genetics 131: 733–740.

Vos, P., R. Hogers, M. Bleeker, M., Reijans, T. Van de Lee, M. Hornes, A. Frijters, J. Pot, J., Peleman, M. Kuiper & M. Zabeau, 1995. AFLP: a new technique for DNA fingerprinting. Nucl Ac Res 23: 4407–4414.

Welsh, W., W. Bushuk, W. Roca & S.P. Singh, 1995. Characterization of agronomic traits and markers of recombinant inbred lines from intra- and interracial populations of *Phaseolus vulgaris* L. Theor Appl Genet 91: 169–177.

Whitlock, M., P. Phillips, F. Moore & S. Tonsor, 1995. Multiple fitness peaks and epistasis. Ann Rev Ecol Syst 26: 601–629.

Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski & S.V. Tingey, 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl Ac Res 18: 6531–6535.

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

Yu, S., J. Li, C. Xu, Y. Tan, Y., Gao, X. Li, Q. Zhang & M. Saghai Maroof, 1997. Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. Proc Natl Acad Sci USA 94: 9226–9231.