Mol Breeding (2010) 25:723–740
DOI 10.1007/s11032-009-9370-z

Protein markers and seed size variation in common bean segregating populations

Ana María González · María De la Fuente · Antonio Miguel De Ron · Marta Santalla

Received: 27 July 2009 / Accepted: 4 December 2009 / Published online: 23 December 2009

Abstract Selection and random genetic drift are the two main forces affecting allele frequencies in common bean breeding programs. Therefore, knowledge on allele frequency changes attributable to these forces is of fundamental importance for breeders. The changes in frequencies of alleles of biochemical markers were examined in F_2 to F_7 populations derived from crosses between cultivated Mesoamerican and Andean common bean accessions (*Phaseolus vulgaris* L.). Biochemical markers included the seed proteins phaseolin, lectin and other seed polypeptides, and six isozymes. The Schaffer's test detected a high significant linear trend of the 63% of the polymorphic loci studied, meaning that directional selection was acting on those loci. Associations between seed size traits, phaseolin seed-storage protein and isozyme markers were detected based on the comparisons of the progeny genotypic means. In the interracial populations the intermediate form PhaH/T, b6, and $Rbcs^{98}$ alleles had a positive effect on seed size. In the inter-gene pool populations, a higher transmission of Mesoamerican alleles in all loci was showed, although the Andean alleles Pha^{T}, $Skdh^{100}$, $Rbcs^{98}$, and $Diap^{100}$ showed positive effects on seed weight. Our results suggest that phaseolin and other seed proteins markers are linked to loci affecting seed size. These markers have good potential for improving the results of the selection and should be considered as a strategy for germplasm enhancement and to avoid the reduced performance of the inter-gene pool populations.

Keywords *Phaseolus vulgaris* L. · Phaseolin · Isozymes · Seed size · Selection · Associations

Introduction

Quantitative traits are often said to be controlled by many genes, each one with an individual effect too small to be measured conveniently under usual experimental conditions (Miles and Wayne 2008). Studies designed to investigate quantitative trait loci (QTL) have proceeded in several ways to overcome this problem. Backcross or F_2 populations have been constructed to maximize linkage disequilibrium between markers and QTL (Nodari et al. 1993; Mutlu et al. 2005). Alternatively, tracking marker frequencies in segregant populations may be used to identify QTL (Edwards et al. 1987). This option has not been utilized as extensively as backcross and F_2 methods, perhaps because of the lack of seed from

Electronic supplementary material The online version of this article (doi:10.1007/s11032-009-9370-z) contains supplementary material, which is available to authorized users.

A. M. González (✉) · M. De la Fuente · A. M. De Ron · M. Santalla
Plant Genetic Resources Department, Misión Biológica de Galicia-CSIC, P.O. Box 28, 36080 Pontevedra, Spain
e-mail: amgonzalez@mbg.cesga.es

cycles of populations under selection, or lack of experimental populations selected for the same trait over many generations. It is likely that genes affecting quantitative traits have a range of effects from large to small, depending on the background in which they are expressed. Theoretically, alleles having positive effects on the trait of interest should increase in frequency with selection and those having negative effects should decrease. The frequency of alleles with large effects should also increase or decrease faster than the frequency of alleles with relatively small effects. However, without easily identifiable alleles, it is difficult to document changes in frequency.

Seed size in cultivated beans has been described as a polygenic trait (Sax 1923) and, from crosses between cultivated and wild types, it has been estimated that there are at least ten genes or blocks of genes involved (Motto et al. 1978). Quantitative inheritance patterns of seed traits such as seed length, width, and height have been reported by Vallejos and Chase (1991a). Nodari et al. (1993) found seven seed weight QTL on linkage groups (LGs) D1, D3, D4, and D7 corresponding to the linkage groups B1, B3, B4, and B7 by Freyre et al. (1998). Of these QTL, those on B1 and B7 correspond to QTL for seed weight also found by Johnson et al. (1996). Park et al. (2000) found five QTL for seed weight (B3, B4, B6, B7, and B8), four QTL for length (three on B8 and one on B2) and three QTL (B4, B6, B11) for height. Guzman-Maldonado et al. (2003) reported five QTL for seed weight in linkage groups B1, B2, B3, and B4. A total of 11 QTL were identified across eight linkage groups by Blair et al. (2006) in an advanced backcross population derived from a cultivated Andean × wild common bean cross. In short, 13 QTL have been confirmed for seed size on the LGs B1, B2, B3, B4, B5, B6, B7, B8, B9, and B11.

Phaseolin and enzymatic markers offer a number of advantages for the analysis of genetic variation which include: (a) lack of detectable effects on the morphology and physiology of the plant, (b) codominant expression, permitting exact identification of genotypes, and (c) lack of epistasis, allowing classification of any number of such markers segregating simultaneously.

Phaseolin and the members of the lectin family (phytohemagglutinin, arcelin and α-amylase inhibitor) belong to vicilins that represent the major seed storage proteins in common bean. A large fraction of these proteins is constituted by phaseolin (Pha) that accounts for 35–46% of total seed proteins in common bean (Mennella et al. 2003). The electrophoretic variability of phaseolin has contributed to the evidence for two gene pools in common bean (Gepts et al. 1986; Singh et al. 1991a; Gepts 1998). The cultivars with "S" and "T" phaseolin patterns predominated in Mesoamerica and in the Southern Andes, respectively. The "B" phaseolin type was presented only in wild and cultivated common beans from Colombia. On the other hand, "C", "H", and "A" phaseolin types were only found among landraces of Andes. Phytohemagglutinin, the major bean seed lectin, accounts for 5–10% of mature seed total protein amount in common bean (Pustzai et al. 1979). More than ten electrophoretic variants of this protein were observed in wild genotypes whereas only a limited heterogeneity was evidenced in domesticated material (Brown et al. 1982; Lioi 1991). Arcelin and α-amylase inhibitor, so far found only in Mesoamerican wild bean accessions, showed seven and four different electrophoretic patterns, respectively (Romero Andreas et al. 1986; Osborn et al. 1988).

Most studies have used isozymes to examine changes in allelic frequencies during selection. Isozymes are useful tools to study the genetic structure of populations, making them good subjects for detailed genetic study. Allelic frequencies at protein loci have been studied in recurrent selection in common bean (Delaney and Bliss 1991), and maize (Stuber et al. 1980; Kahler 1983; Pollak et al. 1984; Butrón et al. 2005). These authors found significant associations between protein frequencies and agronomic traits in selection programs. Alterations of protein frequencies due to selection would cause a skew in the estimation of allele frequencies, as well as a reduction of neutral polymorphism and expected heterozygosity (Kaplan et al. 1989; Stephan et al. 1992). The number of selectively neutral polymorphic sites in a random sample of genes can be reduced by selectively favored substitutions at linked loci. As other consequence, genetic variation that is linked to any beneficial allele will survive the selective phase due to its proximity to the selective target, and the reduction in diversity (measured, e.g., by the number of segregating sites or the average heterozygosity in a sample) would be less. These studies have demonstrated presence of directional

frequency changes in alleles associated with recurrent selection for quantitative traits.

The assessment of the effects of genetic drift and selection is important for designing effective and efficient selection breeding programs. Knowing which marker changes occur in populations during selection might provide information on the genetic components underlying agronomic responses to selection. Further, the use of markers might help to identify specific regions that have been favored by selection. Labate et al. (1999) and De Koeyer et al. (2001) described significant changes in the allele frequencies loci in a long-term recurrent selection programs using RFLP (restriction fragment length polymorphism) markers. Johnson et al. (1996) found a consistent association between the *Phs* locus (phaseolin locus) and seed weight and noted that this association was conserved in other legume species. Weeden and Liang (1985) and Vallejos and Chase (1991a) reported linkage between isozymes and seed components.

The present investigation on characterization of seed size traits and protein markers of several populations developed under selection was undertaken with dual purpose. Firstly determine changes in allele frequencies of protein alleles due to the effects of random genetic drift and selection, and secondly to determine the proportion of the seed size variation that was explained by each of the markers. Assessment of allele frequency changes in breeding programs can be used to detect marker alleles linked to QTL regions under selection pressure.

Materials and methods

Segregant populations

Thirty dry beans inbreeds displaying allelic differences at six enzyme loci, phaseolin and other seed proteins, and difference in seed size were selected from the Misión Biológica de Galicia-Spanish National Research Council (MBG-CSIC) breeding collection (Table 1) to produce F_1 progenies (Table 2).

Part of the seed of each F_1, together with their respective parents, was grown in the field to ascertain hybridity origin and was harvested individually to create F_1-derived F_2 populations. F_3–F_5 populations were planted as single rows and evaluated for seed quality traits (proportion of seed coat, seed size and global quality seed) combined with seed yield. They were grown in plant-to-progeny rows in the F_6 and F_7. F_2 to F_7 seed was stored at 4–5°C of temperature and 50% of humidity. Seeds were carefully dissected in order to germinate and use in the seed storage protein analysis. At the primary leaf stage, tissue samples for isozyme analysis were collected.

Morphological data

Four quantitative seed traits, which defined the seed commercial quality, were scored for each individual: weight (g seed^{-1}), length (mm), defined as the longest distance across the seed parallel to the hilum, height (mm), as the longest distance from the top to the bottom of the seed, and width (mm), measured as the longest distance across the hilum seed.

Seed morphological characters studied reflect the F_{n-1} plant genotype since seed size and seed coat color are maternal characteristics (Roach and Wulff 1987). Therefore, these traits were determined from seeds of the following generation. For example, F_2 phenotypes were determined from $F_{2:3}$ seeds and F_7 phenotypes from $F_{7:8}$ seeds.

Seed storage protein analysis

For storage protein analysis, a portion of the seed was manually removed and ground into a fine dust. The flour sample was suspended at room temperature for at least 0.5 h in a mixture consisting of equal volumes of a 0.5 M NaCl solution (adjusted to pH 2.4 with HCl) and cracking buffer (0.625 M Tris–HCl pH 6.8, 2 mM EDTA; 2% SDS, 40% sucrose, 1% 2-mercaptoethanol and 0.01% bromophenol blue marker dye). Five microliters of sample were subjected to one-dimensional sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS–PAGE) following the method of Brown et al. (1981) modified by Gepts et al. (1986). Electrophoresis was carried out in 1-mm-thick, 15% (wt/vol) polyacrylamide slab gels (height, 16 cm; width, 18 cm) and carried out 25 mA at loading ($\sim$1 h) and thereafter to 30 mA until the separation was completed. Proteins were stained with Coomassie brilliant blue R-250.

Table 1 Characteristics of common bean genotypes used in intra- and interracial hybridization

Line	Race[a]	Seed		Growth habit[b]	Protein[c]				Isoenzymatic alleles[d]					
		Color	Weight (g 100 seed^{-1})		Pha	b4/5	b6/7	b11/12/13	*Skdh*	*Me*	*Rbcs*	*Diap-1*	*Mdh-1*	*Mdh-2*
PHA-020-07	M	White	35	III	S	b5	b7	b13	103	102	98	95	100	100
PHA-119-01	PE	Beige	25	II	H	b4	b7	b12	100	100	100	100	100	100
PHA-159-08	M	White	27	II	S	b5	b7	b13	103	102	98	95	100	100
PHA-159-09	M	White	24	II	S	b5	b7	b13	103	102	98	95	100	100
PHA-159-10	M	White	30	II	S	b5	b7	b13	103	102	98	95	100	100
PHA-159-11	M	White	25	II	B	b5	b7	b11	103	100	98	95	100	100
PHA-159-12	M	White	29	II	B	b5	b7	b11	103	100	100	95	100	102
PHA-159-13	M	White	21	II	S	b5	b7	b13	103	102	100	95	100	100
PHA-159-14	M	White	27	II	S	b5	b7	b13	103	100	100	95	100	100
PHA-257-01	NG	White	70	I	T	b4	b6	b11	100	98	98	100	103	100
PHA-257-04	NG	White	54	I	T	b4	b6	b11	100	98	98	100	103	100
PHA-257-06	NG	White	67	I	T	b4	b6	b11	100	98	98	100	103	100
PHA-257-08	NG	White	58	I	T	b4	b6	b11	100	98	100	100	103	102
PHA-257-10	NG	White	58	I	H	b4	b7	b11	100	98	98	100	103	100
PHA-257-14	NG	White	36	I	T	b4	b6	b12	100	98	98	100	103	100
PHA-257-23	NG	White	63	I	T	b4	b6	b12	100	98	98	100	100	100
PHA-267-15	NG	White	62	I	T	b4	b6	b12	100	100	98	100	103	102
PHA-267-18	NG	White	63	I	T	b4	b6	b11	100	100	98	100	103	100
PHA-267-20	NG	White	51	I	T	b4	b6	b12	100	100	100	100	103	100
PHA-269-12	NG	White	51	I	T	b4	b6	b12	100	98	100	100	100	100
PHA-269-13	NG	White	58	I	T	b4	b6	b12	100	98	100	100	103	100
PHA-272-01	PE	Purple	45	IV	H	b4	b7	b12	100	98	100	100	103	100
PHA-272-02	PE	Purple	43	IV	H	b4	b7	b12	100	98	100	100	103	100
PHA-306-01	PE	Beige	55	I	H	b4	b7	b12	100	100	100	100	100	100
PHA-306-11	PE	Beige	65	I	H	b4	b7	b12	103	100	100	100	100	100
PHA-306-21	PE	Beige	56	I	H	b4	b7	b12	100	100	98	100	103	102
PHA-323-02	CH	Beige	54	IV	C	b4	b6	b13	100	98	100	100	100	100
PHA-338-19	PE	Beige	56	I	H	b4	b7	b12	100	100	98	100	100	100
PHA-338-27	PE	Beige	60	I	H	b4	b7	b12	100	98	100	100	100	100
PHA-452-01	NG	White	74	IV	T	b4	b6	b12	100	98	98	95	103	100

[a] *M* Mesoamerica; *PE* Peru; *NG* Nueva Granada; *CH* Chile

[b] I: Determinate, erect; II: indeterminate, erect; III: indeterminate, non climber or semi-climbing, prostrate; and IV: indeterminate, climbing

[c] Pha, phaseolin; b4/b5, legumin; b6/b7, polypeptides; b11/b12/b13, lectin

[d] *Skdh*, Shikimic acid dehydrogenase; *Me*, malic enzyme; *Rbc*, ribulose bisphosphate carboxylase; *Diap-*, diaphorase; *Mdh-*, malate dehydrogenase

The storage-seed proteins studied were: S, T, C, B, and H phaseolin patterns (43–53 kDa), which were determined according to the genotypes of reference (Sanilac, Tendergreen, Contender, Boyaca and Huevo de Huanchaco). We also present the results of the study of a subset of polypeptides: b4 and b5 (58–55 kDa), b6 and b7 (42–41 kDa), and b11, b12, and b13 polypeptides or lectins (33–31 kDa) (González et al. 2009).

Table 2 Parental and F_7 means for seed traits and the probability of fitting a normal distribution by using the Smirnov–Kolmogorov (*K*) test in the 16 intraracial, interracial and inter-gene pool common bean populations studied

Population P_1/P_2	N^a	Seed															
		Weight				Length				Width				Height			
		P_1	P_2	F_7	*K*	P_1	P_2	F_7	*K*	P_1	P_2	F_7	*K*	P_1	P_2	F_7	*K*
Intraracial populations																	
PHA-20-07/159-11	300	0.25	0.35	0.30	**	12.0	10.1	10.7		8.0	7.4	7.3	**	5.9	5.1	5.3	**
PHA-267-20/257-23	714	0.51	0.63	0.61	**	15.1	16.3	15.5	**	7.5	8.0	8.1	**	6.1	6.7	6.4	**
PHA-338-19/306-21	320	0.56	0.56	0.57	*	14.3	14.0	14.2	**	8.4	8.7	8.4	**	7.2	7.1	7.0	**
Interracial populations																	
PHA-452-01/119-01	50	0.75	0.25	0.62	**	19.7	10.1	16.0		8.7	7.4	7.7		6.4	5.1	6.4	**
PHA-257-01/323-02	125	0.70	0.54	0.80	**	17.2	14.3	17.4	*	8.3	8.8	9.1	**	7.1	7.6	7.4	**
PHA-257-06/306-01	589	0.58	0.65	0.69	**	13.3	13.1	15.4	**	5.7	6.7	8.2	**	4.3	5.1	6.5	**
PHA-257-10/306-11	459	0.67	0.55	0.65	**	16.6	14.6	16.5	**	8.2	8.2	8.3	**	6.8	7.3	6.8	**
PHA-267-18/338-27	515	0.46	0.60	0.66	**	15.3	14.7	16.0		7.5	8.5	8.6	**	6.0	7.1	7.1	**
PHA-272-01/257-01	69	0.45	0.64	0.61	**	12.4	16.5	15.0		8.0	8.1	8.2		6.6	6.5	6.7	
PHA-272-02/257-04	45	0.43	0.54	0.37	**	12.1	15.7	11.9	**	8.2	7.6	8.0		6.4	6.1	5.5	**
Inter-gene pool populations																	
PHA-159-09/257-08	140	0.24	0.58	0.36	**	10.2	15.6	11.2	**	6.9	7.8	7.4		4.5	6.4	5.7	**
PHA-159-12/267-15	148	0.21	0.63	0.41	**	16.1	9.7	13.3	**	7.7	6.6	7.6		6.7	4.5	5.8	**
PHA-159-13/267-18	250	0.27	0.36	0.28	**	10.1	15.3	13.4	**	7.1	7.5	8.2	**	5.3	6.0	6.0	**
PHA-159-14/257-14	150	0.29	0.62	0.27	**	11.8	17.9	11.2	**	7.1	8.4	6.7	*	5.9	6.4	4.9	**
PHA-159-10/269-12	30	0.30	0.51	0.34	**	12.0	15.0	10.2	**	7.2	8.0	7.0	**	5.9	5.9	6.1	
PHA-159-08/269-13	60	0.27	0.58	0.37	**	11.7	15.9	11.3	**	7.1	8.4	8.2	**	5.2	7.0	6.6	**

[a] Number of individuals analyzed in F_7

*, ** Significant at $P \leq 0.05$ or 0.01, respectively

Isozyme analysis

A combination of starch gel electrophoresis and enzyme activity staining was used to detect polymorphisms between parental lines for six enzymes: *Skdh* (E.C.1.1.1.25), *Me* (E.C.1.1.1.40), *Rbcs* (E.C.1.1.1.39), *Prx* (E.C.1.11.1.7), *Mdh-1* (E.C.1.1.1.37), and *Diap* (E.C.1.6.99). Both, *Mdh* and *Diap* enzyme systems had two independent loci.

The seeds used for seed protein analysis were sown, and plant and root tissues were analyzed at the first true-leaf stage, ~10 days after sowing. A crude tissue homogenate was produced by grinding the primary leaf or root apex (depending on the enzymes assayed) in a potassium phosphate grinding buffer (0.1 M pH 7.0 containing 20% sucrose, 5% PVP-40, 0.5%, Triton X-100 and 14 mM of 2-mercaptoethanol). A lithium borate/tris citrate discontinuous system was used. Electrophoresis was carried out at 25 mA for 20 min to load the proteins into the gel, and resumed at 30 mM. After electrophoresis, both anodal and cathodal sections of a gel slice 1.5 mm thick were placed in a tray along with the enzyme assayed, and stained for enzyme activity. Loci were labelled sequentially, with those migrating closest to the anodal end being designated as number 1 (Koenig and Gepts 1989a). In each gel, the cultivars ICA-Pijao and Dark Red Kidney were included as standards. ICA-Pijao has the following genotype at the polymorphic enzyme loci: $Skdh^{103}$, Me^{100}, $Rbcs^{100}$, Prx^{98}, $Diap\text{-}1^{95}$, $Diap\text{-}2^{105}$, $Mdh\text{-}1^{100}$, and $Mdh\text{-}2^{100}$. Dark Red Kidney exhibits the following genotype: $Skdh^{100}$, Me^{98}, $Rbcs^{98}$, Prx^{98}, $Diap\text{-}1^{100}$, $Diap\text{-}2^{100}$, $Mdh\text{-}1^{103}$, and $Mdh\text{-}2^{102}$.

Data analysis

Comparisons of marker frequencies among cycles within populations were made to determine whether significant differences resulted from the selection

process. A regression analysis (SAS 2002) described by Schaffer et al. (1977) and Wilson (1980) was used to determine if the changes in marker frequencies were due to genetic drift or selection. This test relies on the fact that if directional selection is acting on the locus studied, there should be an approximately linear directional trend in the gene frequency data that can be tested for significant deviations from the null hypothesis, i.e., genetic drift acting alone. The variance–covariance matrix for fluctuation in gene frequency was derived for data subjected to the angular transformations 2 $\sin^{-1}$ (marker frequency)$^{1/2}$, as described by Schaffer et al. (1977). The sum of squares of the deviation calculated to test the null hypothesis (changes due to random genetic drift) had a central χ^2 distribution with one degree of freedom less than the number of cycles evaluated. In addition, the sum of squares due to all departures from the model based solely on genetic drift is partitioned in two components, one of which has a central χ^2 distribution with one degree of freedom that accounts for any linear trend in the gene frequency. Therefore, the significance of this linear component determines whether the allelic frequency changes due to directional selection were important. The tabulated χ^2 value (2 *df*, $P \leq 0.01$) critical for declaring a significant deviation was 9.21, and the critical 2 value (1 *df*, $P \leq 0.01$) for declaring a significant linear trend was 6.63.

The proportion of the phenotypic variation that was explained by each one of the protein markers alleles was determined by calculating the coefficients of determination (R^2) due to regression of phenotypic values on allele frequencies in the populations (Edwards et al. 1987). A significant *F*-test indicated cosegregation of the marker locus genotypic classes with the quantitative trait phenotype. For the simplest case of digenic epistasis affecting a quantitative trait, a linear model in a two-way ANOVA (Edwards et al. 1987; Paterson et al. 1988; Devicente and Tanksley 1993; Damerval et al. 1994) was used. Two-factor analyse of variance was computed for each possible pair of loci to determine main effects of the two loci plus their interaction, and the model is as follows:

$$Y_{ijm} = \mu + \alpha_i + \alpha_j + \tau_{ij} + \epsilon_{ijm}, \text{ for } m = 1, 2, \ldots n_{ij},$$

where Y_{ijm} is the phenotype of the mth individuals with the digenic genotype at loci i and j; α_i and α_j are the main effects associated with the loci i and j, respectively; τ_{ij} are the effects arising from interactions between alleles at the loci i and j, and ϵ_{ijm} is the residual effect including the genetic effect unexplained by the two loci in the model plus measurement of errors and other factors, which is assumed to be an independent random variable having a normal distribution.

"Partial R^2" parameters were estimated for each interaction effect and referred to the amount of variation explained by the interaction effect after accounting for the main effects of loci. This was calculated by dividing the Type III sums of squares for the interaction effect by the total sums of squares for the model which included both main effects and the interaction. A significant interaction factor suggests that the effect rendered by the interaction is not simply the sum of their independent effects (Tanksley 1993). In addition, the effect of allelic substitution was estimated as the difference between mean values of the two genotypic classes. These values are relative to the phenotypic means used in the analyses.

Results

Seed traits data

Seed size characteristics exhibited quantitative inheritance patterns with a continuous distribution of phenotypic values in segregation populations. Combining acceptable seed size traits is crucial in the identification of useful genotypes in a breeding program. The selected lines from intra gene pool crosses exhibited higher seed size as compared to the parental lines. The means for seed weight, length, height and width of the parents and F_7 are shown in Table 2. On the one hand, PHA-257-01/323-02 had the highest mean weight, followed by PHA-257-06/306-01, belonging both to interracial populations. In general, seed weight was excellent in most interracial populations. On the other hand, PHA-159-14/257-14 (inter-gene pool population) had the lowest mean, even lower than the intraracial PHA-20-07/159-11 (Mesoamerica × Mesoamerica).

The data did not fit a normal distribution in any population for seed weight, while seed width fitted in five of the 16 populations, seed length in four and seed height in two.

Changes in frequency of seed marker alleles

A fixation (frequency = 1) and an extinction (frequency = 0) of the parental alleles was observed at some of the marker loci. Of the 96 comparisons of marker frequency among generations, 67 significant marker frequency changes were detected in the test for selection (Tables 3, 4).

The effect of selection on changing the frequency of the seed phaseolin alleles differed, depending on the allele and the genetic background. For all populations, the frequencies for the polymorphic markers from F_3 to the final selection cycles are shown in supplementary tables (ESM1, ESM2, and ESM3).

In intra-gene pool populations the phaseolin alleles contributed in a similar way to increase seed weight (Pha^B and Pha^S inside Mesoamerican pool and Pha^T, Pha^H, and Pha^C inside Andean pool), and therefore the most changes in frequency of Pha appeared to be due to random drift.

The frequency of the Pha^T allele increased when it cosegregated with Pha^C allele in one interracial population, while a reduction or stability appeared when it cosegregated with Pha^H in the rest of interracial populations or with Pha^S in inter-gene pool populations. The Pha^T frequency changes exhibited significance in the test for selection only in two of the five crosses in which segregated with Pha^H (Table 3). Moreover, this allele showed non-linearity when it segregated with Pha^C. Thus the changes in frequency of Pha^T appeared to be due to random drift in the interracial populations, indicating similar contribution of Pha^T, Pha^H, and Pha^C in seed weight. However, there was a highly linear trend in the frequencies of Pha^T allele in the inter-gene pool populations (Table 4).

Pha^S allele did not exhibit significant linear increases in frequency in the intraracial population PHA-20-07/PHA-159-11. In the inter-gene pool populations, Pha^S increased or kept stable throughout the generations. The Schaffer test detected that the allelic frequency changes of Pha^S were significantly greater than those expected from random genetic drift acting alone in the inter-gene pool crosses (Table 4). The χ^2 of the directional selection effect (χ^2 lineal) was highly significant for this allele in all populations. Our finding suggests that selection was not effective in increasing seed size in inter-gene pool

Table 3 Chi square values for deviations from genetic drift and for directional selection of significant allelic frequency changes at protein markers in intra-gene pool crosses

Marker	AFCC[a]	χ^2 drift	χ^2 selection
Intraracial crosses			
PHA-267-20/PHA-257-23[b]			
Rbc^{100}/Rbc^{98}	0.09	160.9***	22.70***
$Mdh1^{103}/Mdh1^{100}$	0.12	70.38***	16.14***
PHA-338-19/PHA-306-21[b]			
$Mdh2^{100}/Mdh2^{102}$	0.07	167.81***	61.22***
Interracial crosses			
PHA-452-01/PHA-119-01[b]			
Phaseolin H/T	0.04	7.28	16.49**
b6/b7	0.17	37.49**	32.95***
PHA-257-01/PHA-323-02[b]			
b11/b13	0.38	378.8***	204.49***
Rbc^{100}/Rbc^{98}	0.10	1.18	42.96***
$Mdh1^{103}/Mdh1^{100}$	0.10	505.05***	215.6***
PHA-257-06/PHA-306-01[b]			
Phaseolin H/T	0.04	128.82***	105.21**
b6/b7	0.09	697.74***	96.16***
b11/b12	0.07	61.15***	13.43***
Me^{100}/Me^{98}	0.19	86.39***	15.94***
Rbc^{100}/Rbc^{98}	0.09	12.67	8.86***
$Mdh1^{103}/Mdh1^{100}$	0.19	71.21***	17.19***
PHA-257-10/PHA-306-11[b]			
$Skdh^{103}/Skdh^{100}$	0.08	34.74***	11.28***
Me^{100}/Me^{98}	0.14	40.45***	22.19***
Rbc^{100}/Rbc^{98}	0.24	49.46***	28.38***
$Mdh1^{103}/Mdh1^{100}$	0.14	30.02***	13.32***
PHA-267-18/PHA-338-27[b]			
b11/b12	0.11	52.34***	7.41**
Me^{100}/Me^{98}	0.06	7.87	5.15**
Rbc^{100}/Rbc^{98}	0.06	199.86***	122.72***
$Mdh1^{103}/Mdh1^{100}$	0.11	41.52***	22.81***
PHA-272-01/PHA-257-01[b]			
b6/b7	0.15	7.85	7.72**
Rbc^{100}/Rbc^{98}	0.06	3.75	3.32***
PHA-272-02/PHA-257-04[b]			
b6/b7	0.17	13.76***	7.43**
Rbc^{100}/Rbc^{98}	0.12	10.36	7.70**

[a] Average frequency change per cycle

[b] Race Nueva Granada = PHA-267-20, PHA-257-23, PHA-452-01, PHA-257-01, PHA-257-04, PHA-257-06, PHA-257-10, and PHA-267-18. Race Peru = PHA-338-19, PHA-306-21, PHA-306-01, PHA-306-11, PHA-338-27, PHA-272-01, PHA-272-02, PHA-119-01. Race Chile = PHA-323-02

, * Significant at $P \leq 0.01$ or 0.005, respectively

Table 4 Chi square values for deviations from genetic drift and for directional selection of significant allelic frequency changes at protein markers in the inter-gene pool crosses

Marker	AFCC[a]	χ^2 drift	χ^2 selection
PHA-159-09/PHA-257-08[b]			
Phaseolin S/T	0.025	0.99	20.71***
b6/b7	0.182	6.27	10.58***
b11/b13	0.483	66.29***	23.88***
$Skdh^{103}/Skdh^{100}$	0.095	7.05	32.22***
Me^{102}/Me^{98}	0.102	26.89***	25.89***
Rbc^{100}/Rbc^{98}	0.244	104.66***	19.53***
$Diap1^{100}/Diap1^{95}$	0.030	347.86***	175.28***
$Mdh1^{103}/Mdh1^{100}$	0.090	16.61***	14.92***
PHA-159-12/PHA-267-15[b]			
Phaseolin B/T	0.245	11.80***	34.60***
b6/b7	0.245	26.02***	26.00***
$Skdh^{103}/Skdh^{100}$	0.249	52.81***	25.60***
Rbc^{100}/Rbc^{98}	0.189	66.80***	62.72***
$Diap1^{100}/Diap1^{95}$	0.089	121.14***	96.09***
$Mdh1^{103}/Mdh1^{100}$	0.325	333.40***	212.52***
PHA-159-13/PHA-267-18[b]			
Phaseolin S/T	0.157	11.18***	61.15***
b4/b5	0.153	57.91***	33.30***
b6/b7	0.136	81.72***	39.60***
b11/b13	0.169	51.09***	23.62***
$Skdh^{103}/Skdh^{100}$	0.145	54.60***	21.41***
Rbc^{100}/Rbc^{98}	0.059	395.96***	199.01***
$Diap1^{100}/Diap1^{95}$	0.080	122.95***	42.11***
$Mdh1^{103}/Mdh1^{100}$	0.093	936.27***	342.97***
PHA-159-14/PHA-257-14[b]			
Phaseolin S/T	0.045	12.52***	22.32***
b12/b13	0.282	201.96***	75.16***
$Skdh^{103}/Skdh^{100}$	0.278	245.26***	77.07***
Me^{100}/Me^{98}	0.191	81.57***	32.53***
$Mdh1^{103}/Mdh1^{100}$	0.200	233.84***	155.36***
PHA-159-10/PHA-269-12[b]			
Phaseolin S/T	0.167	2.34	8.23***
b12/b13	0.095	91.85***	57.20***
$Skdh^{103}/Skdh^{100}$	0.258	39.29***	25.64***
Me^{100}/Me^{98}	0.130	54.35***	19.89***
$Diap1^{100}/Diap1^{95}$	0.206	125.06***	71.28***
PHA-159-08/PHA-269-13[b]			
Phaseolin S/T	0.109	5.09	85.67***
b4/b5	0.398	26.18***	9.39***
b6/b7	0.0220	35.70***	16.28***
b12/b13	0.112	113.04***	73.87***
$Skdh^{103}/Skdh^{100}$	0.136	29.10***	15.82***
Me^{102}/Me^{98}	0.121	714.85***	603.24***
Rbc^{100}/Rbc^{98}	0.243	23.62***	8.31***
$Diap1^{100}/Diap1^{95}$	0.231	106.33***	72.86***
$Mdh1^{103}/Mdh1^{100}$	0.400	125.27***	104.76***

[a] Average frequency change per cycle

[b] Race Mesoamerica = PHA-159-08, PHA-159-09, PHA-159-10, PHA-159-12, PHA-159-13, and PHA-159-14. Race Nueva Granada = PHA-257-08, PHA-267-15, PHA-267-18, PHA-257-14, PHA-269-12, and PHA-269-13

, * Significant at $P \leq 0.01$ or 0.005, respectively

crosses in spite of differences in protein frequencies were observed. These changes were due to an increase of the S haplotype that exhibited dominance over the T haplotype and gave rise to lines with a smaller seed size than parentals.

Two of the isozyme loci examined (Prx^{98} and $Diap\text{-}2^{105}$) did not revealed polymorphisms in any population studied. The alleles that encode for the remaining isoenzymes suffered a selective pressure in the crosses analyzed, increasing the frequencies of the alleles Me^{98}, $Rbcs^{98}$, and $Mdh1^{103}$ (Andean alleles) in the interracial populations, while in the inter-gene pool populations showed an increase of the Mesoamerican alleles $Skdh^{103}$, $Rbcs^{100}$, and $Diap1^{95}$.

The selection procedure resulted in a comparatively high selection response in interracial populations, which could be due to the accumulation of large-seeded alleles. Inter-gene pool populations showed an overrepresentation of the Mesoamerican alleles, many of which were fixed in the F_7 lines. The presence of the great percentage of Mesoamerican markers and the lowest mean in these populations revealed a strong association between seed size and the markers studied. These results revealed a great chance to find new favourable combinations of alleles and they could indicate a selective advantage on genetic level.

Proportion of variation explained by protein alleles

The genetic control and the location of phaseolin (Pha), lectins (*b11/b12/b13*), *Rbcs*, *Skdh*, *Me*, *Diap*, and *Mdh* has been previously described (Vallejos and Chase 1991a, b). These markers are distributed on

seven linkage groups. The markers *b4/b5* and *b6/b7* have not been described previously in common bean.

The polypeptides of 56 or 54 kDa (*b4/b5*) and 42 or 41 kDa (*b6/b7*) molecular weights segregated as products of different alleles of the same gene. In the classification of parental lines screened (Table 1), all lines with S-type phaseolin had the 54 kDa (*b5*) and 41 kDa (*b7*) polypeptides, and most lines with T, H or C type phaseolin had the 56 and 42 kDa (*b4* and *b6*) bands. Gepts et al. (1986) demonstrated that bean cultivars and landraces with S type phaseolin were domesticated in Mexico and Central America, and that beans with T- and C-type phaseolin were domesticated in the Andes. Although the number of lines screened is only a small sample, it suggests the possibility that the polypeptide patterns are also indicative of both centers of domestication.

In order to investigate possible linkage between seed size and protein and isozyme loci, *F*-test was used to compare the means of seed size parameters between homozygotes and heterozygotes at all segregating marker loci. The means of seed weight, length, height, and width were grouped according to class frequencies for each marker: protein band and isoenzyme. Those showing significant ($P \leq 0.05$) differences between the groups mean values for these traits are shown in Tables 5 and 6. Inspection of the mean seed weights of the three genotypes indicated that the mode of action of these genes are additive. The mean seed weight of heterozygotes was intermediate to the two homozygote classes.

Phaseolin explained from 5 to 26% of the phenotypic variance of seed weight. Substitution of a lower-seed weight allele by a higher-seed weight allele resulted in a seed weight from 1 to 25 g 100 seed^{-1}, depending on the genetic background. For the *b6/b7* and *b4/b5* polypeptides, R^2 values ranged from 7 to 20% and from 6 to 24%, respectively. Thus, both markers, *b6/b7* and *b4/b5*, could be also useful for studying seed weight variability. Significant associations between isozymes markers and seed weight explained notable ranges of phenotypic variation: *Skdh* (4–14%), *Rbcs* (5–12%), and *Diap-1* (4–10%).

Phaseolin explained from 7 to 48% of the phenotypic variance of seed length, which was larger than the coefficient of determination (R^2) explained for seed weight. The polypeptides *b6/b7* and *b4/b5* explained from 0.8 to 16% and from 6 to 8% of the phenotypic variance, respectively.

The lectins explained from 4 to 16% of the phenotypic variation of seed height. The interaction effect of loci of lectins and *Me* ranged from 4 to 7% of the phenotypic variation after accounting for the main effects of the loci, while only one population showed interaction effect of loci of lectins with *Rbcs* (3%).

Mdh-1 explained the largest percentage of phenotypic variation for seed width, from 6 to 12%. The polypeptide *b6/b7* explained 4–10% of the phenotypic variance. Both markers showed significant associations and explained from 3 to 6% of the phenotypic variance.

Vallejos and Chase (1991b) mapped five seed proteins, the *P* locus, and *Mdh-1* in the linkage group (B7) of the Pha locus. A significant interaction was observed between the polypeptides *b4/b5* and *b6/b7* and the isozyme *Mdh-1* in several crosses, indicating that these loci characterize close regions of the bean genome. These polypeptides must be positioned on the LG B7, closer than *Mdh-1* than Pha.

In the interracial populations the intermediate form PhaH/T, the homozygous b6 and $Rbcs^{98}$ allele had a positive effect on seed size. The inter-gene pool populations increased the number of potential markers, showing an effect positive on seed size PhaT, b4, b6, $Skdh^{100}$, $Rbcs^{98}$, and $Diap^{100}$ alleles.

The observed allelic effects are shown in Table 7. An allelic effect of −0.5 for seed size traits indicates that the parental allele reduced seed size by half a percentage point in that cross. Although allelic effects varied depending on environmental conditions and genetic backgrounds, it has been noted that Andean alleles contributed a large seed size.

Discussion

The multigenic type of inheritance is in accord with the studies related to seed size in *Phaseolus* (Motto et al. 1978; Nienhuis and Singh 1988; Vallejos and Chase 1991a; Park et al. 2000). The highest means and the widest variation for seed size traits belonged to interracial populations. This demonstrated the great usefulness of the interracial populations for improvement of large-seeded germplasm (Singh and Urrea 1994; Welsh et al. 1995; Singh et al. 2002). However, the seed size improvement within intraracial and inter-gene populations was very poor. One

Table 5 Coefficient of determination (R^2), significant means of marker classes and partial R^2 for seed weight and seed length of the 16 intraracial, interracial and inter-gene pool common bean populations

Marker/interaction	LG[a]	Seed weight (g 100 seed^{-1})				Seed length (mm)			
		R^{2b}	MM[c]	MS	SS	R^2	MM	MS	SS
Intraracial crosses									
PHA-20-07/PHA-159-11									
Pha B/S	B7	5.1	0.29	0.31	0.33***	9.1	10.7	10.2	11.0***
Me$^{102/100}$	B4	3.5	0.32	0.34	0.31***				
*b11/b13*Me*$^{102/100}$		3.1			***	3.6			**
PHA-267-20/PHA-257-23									
Me$^{100/98}$	B4	1.1	0.53	0.67	0.56*				
Rbcs$^{100/98}$	B4	8.6	0.48	0.64	0.58***				
Interracial crosses									
PHA-452-01/PHA-119-01									
Pha H/T	B7	6.1	0.59	0.73	0.70**	11.4	14.2	17.6	17.2***
b6/b7		7.2	0.72	0.57	0.49**	6.3	17.4	15.1	15.4***
*b6/b7*Mdh-1*$^{103/100}$		5.1			**				
PHA-257-01/PHA-323-02									
Pha T/C	B7					7.0	16.8	15.5	15.7***
Rbcs$^{100/98}$	B4	8.8	0.60	0.70	0.75***				
*b11/b13*Rbcs*$^{100/98}$		3.1							
PHA-257-06/PHA-306-01									
Pha H/T	B7	19.1	0.71	0.96	0.83***	9.0	16.0	17.5	16.7***
b6/b7		11.8	0.90	0.83	0.67***	4.4	18.6	17.5	16.7***
Me$^{100/98}$	B4	4.2	0.63	0.82	0.67***				
Rbcs$^{100/98}$	B4	4.6	0.65	0.68	0.74**				
Pha H/T**Mdh-1*$^{103/100}$		2.0			*				
*b6/b7*Mdh-1*$^{103/100}$		4.3			***				
PHA-257-10/PHA-306-11									
Skdh$^{103/100}$	B3	4.2	0.42	0.55	0.53**				
Rbcs$^{100/98}$	B4	11.7	0.41	0.83	0.63**				
PHA-267-18/PHA-338-27									
Pha H/T	B7	12.9	0.68	0.59	0.66***	12.0	15.0	16.2	15.2**
b6/b7		10.4	0.69	0.65	0.52***				
Rbcs$^{100/98}$	B4	4.9	0.60	0.64	0.69*				
Diap-1$^{100/95}$	B5	4.3	0.65	0.62	0.59***				
Me$^{100/9}$**b11/b12*	B4	6.7			***				
PHA-272-01/PHA-257-01									
Pha H/T	B7	13.5	0.54	0.69	0.64***	16.3	14.5	16.0	15.3***
b6/b7		16.1	0.70	0.65	0.57***	15.2	15.8	15.1	14.5***
Pha H/T**b6/b7*		6.1			***				
PHA-272-02/PHA-257-04									
Pha H/T	B7	21.1	0.41	0.61	0.58***	18.8	12.6	15.6	14.3***
b6/b7		19.6	0.80	0.74	0.51***	15.4	14.7	14.5	12.6***
b11/b12	B4	8.2	0.72	0.83	0.78***				
Pha H/T**b6/b7*		6.0			***				

Table 5 continued

Marker/interaction	LG[a]	Seed weight (g 100 seed^{-1})				Seed length (mm)			
		R^{2b}	MM[c]	MS	SS	R^2	MM	MS	SS
Inter-gene pool crosses									
PHA-159-09/PHA-257-08									
b4/b5		5.9	0.46	0.43	0.36***				
Pha S/T	B7	10.9	0.31	0.39	0.65***	22.0	12.1	12.8	18.1***
Skdh$^{103/100}$	B3	5.6	0.36	0.40	0.44***				
Me$^{102/98}$	B4	2.2	0.36	0.43	0.41*				
Rbcs$^{100/98}$	B4	9.6	0.33	0.41	0.46***				
Diap-1$^{100/95}$	B5	10.2	0.43	0.37	0.36***				
*b4/b5*Mdh-1*$^{103/100}$		3.8			**				
*b6/b7*Mdh-1*$^{103/100}$		5.2			***				
PHA-159-12/PHA-267-15									
b4/b5		24.0	0.45	0.29	0.32***				
Pha B/T	B7	25.6	0.31	0.42	0.47***	48.2	10.9	13.6	14.3***
b6/b7		15.8	0.47	0.31	0.34***	15.1	14.3	10.5	11.5***
Rbcs$^{100/98}$		11.6	0.32	0.28	0.44***	9.6	11.2	10.1	13.7***
PHA-159-13/PHA-267-18									
b4/b5		10.5	0.48	0.45	0.36***	4.0	13.1	13.5	12.6***
Pha S/T	B7	13.0	0.30	0.39	0.44***				
b6/b7		15.9	0.49	0.40	0.34***	0.8	13.7	13.1	13.0*
Skdh$^{103/100}$	B3	5.8	0.40	0.45	0.48***				
Rbcs$^{100/98}$	B4	6.3	0.40	0.42	0.47***				
Diap-1$^{100/95}$	B5	8.5	0.50	0.47	0.40***				
*b4/b5*b6/b7*						7.4			***
*b4/b5*Mdh-1*$^{103/100}$		1.8			*				
Pha S/T**Mdh-1*$^{103/100}$		1.6			*				
*b6/b7*Mdh-1*$^{103/100}$		2.6			**				
PHA-159-14/PHA-257-14									
Pha S/T	B7	15.2	0.25	0.30	0.57***	24.6	10.6	11.7	15.5***
b6/b7		10.2	0.35	0.31	0.27***	2.2	11.8	11.3	11.8***
Skdh$^{103/100}$	B3	12.3	0.27	0.34	0.40***				
Diap-1$^{100/95}$	B5	4.7	0.41	0.32	0.24***				
PHA-159-10/PHA-269-12									
b4/b5		12.5	0.40	0.35	0.31***	8.4	12.8	11.5	10.2***
Pha S/T	B7	23.9	0.78	0.80	0.83***	22.2	10.6	11.1	13.3***
b6/b7		14.3	0.41	0.36	0.31***	5.9	12.6	12.1	10.8***
Skdh$^{103/100}$	B3	14.3	0.31	0.27	0.37***				
PHA-159-08/PHA-269-13									
b4/b5		10.9	0.47	0.37	0.37***	3.1	12.8	12.4	12.2***
Pha S/T	B7	12.0	0.32	0.42	0.57**	17.8	11.3	12.6	15.9***
b6/b7		14.4	0.54	0.48	0.41***	5.9	14.2	13.6	12.5**
Skdh$^{103/100}$	B3	5.6	0.39	0.35	0.44**				
*b4/b5*b6/b7*						6.1			**
*b4/b5*Mdh*$^{103/100}$		3.8			**				

Table 5 continued

Marker/interaction	LG[a]	Seed weight (g 100 seed^{-1})				Seed length (mm)			
		R^{2b}	MM[c]	MS	SS	R^2	MM	MS	SS
Pha S/T*$Mdh^{103/100}$		2.2			*				
*b6/b7**$Mdh^{103/100}$		5.7			***				

[a] Linkage group

[b] Coefficient of determination: phenotypic variation explained by the marker

[c] MM and SS are homozygous for parental alleles; MS is heterozygous

*, **, *** Significant at $P \leq 0.05$, 0.01 or 0.001, respectively

explanation for intraracial populations is the narrow genetic base that restricts the genetic gain. Possible causes for inter-gene populations include genetic factors affecting the transmission of the genes (González et al. 2009).

The data indicated that seed width showed a normal distribution in most populations. This could be due to the fact that width is a more stable trait, and it is not so important to identify common bean market classes, compared to seed weight, so it has gone more "unnoticed" to the selection process used in this study.

Selection increases the frequencies of favourable alleles while genetic drift is a random change in allele frequencies due to small population size. A loss of favourable alleles due to random genetic drift leads to a reduction in genetic variance and, thus, limits future selection response (Guzman and Lamkey 2000). Although the loss of genetic variation was generally consistent with that expected for a model in which random genetic drift acts alone on neutral alleles, the changes observed in the frequency of a large part of the markers were significantly greater from that which might be expected under such conditions. The selection resulted in the decrease of the genetic variation, thus random fluctuations were increased and caused that the genes of one generation were not a representative sample of the previous generation. As a result, a great part of the deviations were due to the action of any erratic factor, such as non-directional selection combined with random genetic drift and/or sampling error.

The Schaffer's test detected a high significant linear trend of the 63% of the polymorphic loci studied, meaning that directional selection was acting on those loci (Schaffer et al. 1977). These results agreed with other works (Eagen and Goldman 1996; Heredia-Diaz et al. 1996; Butrón et al. 2005; Diaby and Casler 2005), which found significant changes in marker frequencies due to directional selection. The higher power of this study for detecting directional selection effects on allelic frequencies could be due to that in those studies the cycles evaluated were the result of selecting for different traits. Some of them are complex traits, as grain yield, and other traits could be negatively correlated, so the selection to increase one trait can result in the reduction of other trait, while in this study we selected for seed size, seed weight, length, height and width that are traits positively correlated (Vallejos and Chase 1991a; Park et al. 2000).

The difference in seed size between the progenitors suggests that either a large number of polygenes with similar effects or at least one major and a few others with minor effects would segregate in the progenies. With a limited number of markers, one can expect to detect minor genes with tight linkages or major genes with moderate linkage. The consistent results obtained with 16 independently obtained crosses strengthened the validity of the linkage between protein markers and the locus that contributes to seed size.

The statistically significant associations possibly infer a genetic linkage between the markers and the seed size traits in form of multilocus associations and/or pleiotropic effects of genes controlling these traits. Associations were found between T/H phaseolin protein and large seed size, and S phaseolin and small seed size. This association was consistent across the wide range of populations analyzed, and agreed with Johnson et al. (1996) and Murray et al. (2004) who also found a stable relationship between the locus *Phs* and seed weight in different environments. The locus *Phs* was significantly associated with seed weight in all populations in which it was polymorphic, except for the population PHA-257-01/

Table 6 The coefficient of determination (R^2), significant means of marker classes and partial R^2 and for seed height and width of the 16 intraracial, interracial and inter-gene pool common bean populations

Marker/interaction	LG[a]	Seed height (mm)				Seed width (mm)			
		R^{2b}	MM[c]	MS	SS	R^2	MM	MS	SS
Intraracial populations									
PHA-20-07/PHA-159-11									
b11/b13	B4	4.1	7.4	7.7	6.8**				
Me$^{102/100}$	B4	0.9	7.5	7.6	7.4*				
*b11/b13*Me*$^{102/100}$		4.4			***				
PHA-267-20/PHA-257-23									
Me$^{100/98}$	B4	1.2	7.8	8.6	7.8***				
Rbcs$^{100/98}$	B4	6.7	7.5	8.4	7.9***				
PHA-338-19/PHA-306-21									
Mdh-1$^{103/100}$	B7					9.0	7.1		6.8***
Interracial populations									
PHA-257-01/PHA-323-02									
b11/b13	B4	9.1	8.5	8.0	9.9***	2.1	7.0	6.4	7.3**
Rbcs$^{100/98}$	B4	3.2	8.9	9.1	9.8***				
Mdh-1$^{103/100}$	B7					6.9	7.8	7.6	6.1*
*b11/b13*Rbc*$^{100/98}$		3.1			**				
PHA-257-06/PHA-306-01									
Me$^{100/98}$	B4	12.4	8.3	8.3	9.3***				
Mdh-1$^{103/100}$	B7					8.6	6.4	7.3	7.0***
*b11/b12*Me*$^{100/98}$		6.5			**				
PHA-257-10/PHA-306-11									
b11/b12	B4					1.8	6.0	6.2	6.3***
Mdh-1$^{103/100}$	B7					8.9	6.3	5.9	5.8***
PHA-267-18/PHA-338-27									
b6/b7						5.1	6.9	7.2	7.3***
b11/b12	B4	15.9	7.1	8.8	8.9***				
Mdh-1$^{103/100}$	B7					8.5	7.2		6.4***
*b6/b7*Mdh-1*$^{103/100}$						6.4			***
PHA-272-01/PHA-257-01									
b11/b12	B4	14.2	8.1	8.5	8.9***				
PHA-272-02/PHA-257-04									
b6/b7						9.9	6.4	7.0	5.7***
Mdh-1$^{103/100}$	B7					8.3	6.9	6.1	5.8***
*b6/b7*Mdh-1*$^{103/100}$						6.2			*
Inter-gene pool populations									
PHA-159-09/PHA-257-08									
b6/b7						4.1	6.9	6.4	6.2**
b11/b13	B4	8.0	7.6	7.9	7.9*	1.7	6.0	6.1	6.3*
PHA-159-12/PHA-267-15									
b6/b7						4.4	6.0	5.0	5.6***
b11/b13	B4	15.5	7.4	6.9	7.8***				
Rbcs$^{100/98}$	B4	9.0	7.2	7.3	8.5***				

Table 6 continued

Marker/interaction	LG[a]	Seed height (mm)				Seed width (mm)			
		R^{2b}	MM[c]	MS	SS	R^2	MM	MS	SS
Mdh-1$^{103/100}$	B7					12.1	6.3	5.6	5.1***
PHA-159-13/PHA-267-18									
b6/b7						6.8	5.3	5.5	5.1*
b11/b13	B4	14.2	8.0	7.7	8.2***				
Mdh-1$^{103/100}$	B7					12.3	5.9	5.7	5.5**
*b6/b7*Mdh-1*$^{103/100}$				2.5			**		
PHA-159-14/PHA-257-14									
Pha S/T	B7	4.0	6.9	7.1	7.8***				
b6/b7						4.3	5.1	5.3	4.9***
b12/b13	B4	12.1	7.2	7.3	6.6***				
Me$^{100/98}$	B4	15.6	6.8	7.1	7.8**				
Mdh-1$^{103/100}$	B7					6.1	5.2	5.0	4.7**
Me$^{100/98}$**b12/b13*		7.1			***				
PHA-159-10/PHA-269-12									
b6/b7						3.4	5.9	5.7	5.6*
b12/b13	B4	10.0	7.7	8.0	7.2***				
Me$^{100/98}$	B4	14.8	7.1	7.4	8.0***				
PHA-159-08/PHA-269-13									
Me$^{100/98}$	B4	10.5	8.2	8.4	8.8***				
b12/b13	B4	11.3	9.0	8.5	8.3*	2.4	6.7	5.2	6.0***
Mdh-1$^{103/100}$	B7					8.5	5.7	6.5	6.2***

[a] Coefficient of determination: phenotypic variation explained by the marker

[b] MM and SS are homozygous for parental alleles; MS is heterozygous

*, **, *** Significant at $P \leq 0.05$, 0.01 or 0.001, respectively

PHA-323-02, in which the parents corresponded to T and C phaseolin patterns, respectively. Hartana (1983) reported that T and C phaseolin patterns were associated with large seed weight. This study showed that the alleles of both phaseolin patterns (T and C) contributed in a similar way to increase seed weight. The larger magnitude of the phenotypic variance explained by seed length than seed weight implies that the QTL for seed length could be mapped nearer locus *Phs* than the QTL for seed weight.

These populations represent the major variability of both gene pools in common bean, so the locus *Phs* provides an attractive model system for the molecular analysis of quantitative trait variation and the genetic expression of seed size. In addition, due to the large R^2 observed in this study, the locus *Phs* and some polypeptides, as b6/b7, could be useful for MAS (Marker Assisted Selection).

In common bean, associations between *Rbcs* and *Me* have been reported in different populations (Koenig and Gepts 1989b; Nodari et al. 1993). The interaction effect of the lectins (*b11/b12/b13*), *Me* and *Rbcs* confirm the order *Rbcs*-Lec-*Me* reported by Koenig and Gepts (1989b). The lectins should be situated in LG B4, closer to *Me* than *Rbcs*.

The significant interaction observed between the polypeptides *b4/b5* and *b6/b7* and the isozyme *Mdh-1* indicates that these polypeptides should be situated in the same LG B7, near the locus *Phs*.

The isoenzymes are a useful tool, as the phaseolin, to characterize both gene pools, being characteristic the alleles *Skdh*^{103}, *Rbcs*^{100}, *Mdh1*^{100}, and *Diap1*^{95} from the Mesoamerican pool, while the alleles *Skdh*^{100}, *Rbcs*^{98}, *Mdh1*^{103}, and *Diap1*^{100} characterize Andean pool (Singh et al. 1991b). Intra-gene pool populations showed an increase of isoenzymatic

Table 7 Allelic effects for seed size traits of the intraracial, interracial and inter-gene pool common bean populations

Allelic effect[a]				
Marker	% weight (100 seed^{-1})	% length (mm)	% height	% width
Intraracial crosses				
PHA-20-07/PHA-159-11				
Pha B/S	−0.04	−0.3		
b11/b13			0.6	
Me$^{102/100}$	0.01		0.1	
PHA-267-20/PHA-257-23				
Me$^{100/98}$	−0.03			
Rbcs$^{100/98}$	−0.10		−0.4	
Interracial crosses				
PHA-452-01/PHA-119-01				
Pha H/T	−0.11	−3		
b6/b7	0.23	2		
PHA-257-01/PHA-323-02				
Pha T/C		1.1		
b11/b13			−1.4	−0.3
Rbcs$^{100/98}$	−0.15		−0.9	
Mdh-1$^{103/100}$				1.7
PHA-257-06/PHA-306-01				
Pha H/T	−0.12	−0.7		
b6/b7	0.23	1.9		
Me$^{100/98}$	−0.04		−1	
Rbcs$^{100/98}$	−0.09			
Mdh-1$^{103/100}$				−0.6
PHA-257-10/PHA-306-11				
b11/b12				−0.3
Skdh$^{103/100}$	−0.11			
Rbcs$^{100/98}$	−0.22			
Mdh-1$^{103/100}$				0.5
PHA-267-18/PHA-338-27				
Pha H/T	0.02	−0.2		
b6/b7	0.17			−0.4
b11/b12			−1.8	
Rbcs$^{100/98}$	−0.09			
Diap-1$^{100/95}$	0.06			
Mdh-1$^{103/100}$				0.8
PHA-272-01/PHA-257-01				
Pha H/T	−0.10	−0.8		
b6/b7	0.13	1.3		
b11/b12			−0.8	
PHA-272-02/PHA-257-04				
Pha H/T	−0.17	−1.7		
b6/b7	0.29	2.1		0.7
b11/b12	−0.06			
Mdh-1$^{103/100}$				1.1
Inter-gene pool crosses				
PHA-159-09/PHA-257-08				
b4/b5	0.10			
b6/b7				0.7
b11/b13			−0.3	−0.3
Pha S/T	−0.34	−6		
Skdh$^{103/100}$	−0.08			
Me$^{102/98}$	−0.05			
Rbcs$^{100/98}$	−0.13			
Diap-1$^{100/95}$	0.07			
PHA-159-12/PHA-267-15				
b4/b5	0.13			
Pha B/T	−0.16	−3.4		
b6/b7	0.13	2.8		0.4
b11/b13			−0.4	
Rbcs$^{100/98}$	−0.12	−2.5	−1.3	
Mdh-1$^{103/100}$				1.2
PHA-159-13/PHA-267-18				
b4/b5	0.12	0.5		
Pha S/T	−0.14			
b6/b7	0.15	0.7		0.2
b11/b13			−0.2	
Skdh$^{103/100}$	−0.08			
Rbcs$^{100/98}$	−0.07			
Diap-1$^{100/95}$	−0.10			
Mdh-1$^{103/100}$				0.4
PHA-159-14/PHA-257-14				
Pha S/T	−0.32	−4.9	−0.9	
b6/b7	0.08	0		0.2
b12/b13			0.6	
Skdh$^{103/100}$	−0.13			
Me$^{100/98}$			−1	
Diap-1$^{100/95}$	0.17			
Mdh-1$^{103/100}$				0.8
PHA-159-10/PHA-269-12				
b4/b5	0.09	2.6		
Pha S/T	−0.05	−2.7		
b6/b7	0.10	1.8		0.3
b12/b13			0.5	
Skdh$^{103/100}$	−0.06			
Me$^{100/98}$			−0.9	

Table 7 continued

Allelic effect[a]				
Marker	% weight (100 seed^{-1})	% length (mm)	% height	% width
PHA-159-08/PHA-269-13				
b4/b5	0.10	0.6		
Pha S/T	−0.25	−4.6		
b6/b7	0.13	1.7		
b12/b13			0.7	
Skdh$^{103/100}$	−0.05			
Me$^{100/98}$			−0.6	
Mdh-1$^{103/100}$				−0.5

[a] The allelic effect correspond to the first allele of the segregating pair in the population

alleles: *Me*^{98}, *Rbcs*^{98}, and *Mdh1*^{103}. The frequencies of the b6 and *Rbcs*^{98} alleles shifted in a manner consistent with its association with seed size. Both alleles, which have a positive effect on seed size, were increased markedly with each generation. Deviations for directional selection were significant, indicating that the changes were too large to be due to random drift alone.

The inter-gene pool populations showed a higher transmission of Mesoamerican alleles in all loci, although the Andean alleles, as Pha^{T}, *Skdh*^{100}, *Rbcs*^{98}, and *Diap*^{100}, showed positive effects on seed weight. The summary of the allelic effects on seed size within different genetic backgrounds showed a clear tendency for Andean alleles to increase seed size. Johnson et al. (1996) determinate that S haplotype was associated with lower seed weight and exhibited some degree of dominance over the T haplotype. Segregation distortion is a common feature of inter-gene pool crosses in common bean (Freyre et al. 1998; Johnson and Gepts 2002). It has been found that in crosses of Andean × Mesoamerican parents, those from the Mesoamerican gene pool usually possess negative general combining abilities for seed size (Singh 1992). The negative phenotypic effect from Mesoamerican alleles was reflected in reduced performance of these populations. The bad performance of the progenies of Andean × Mesoamerican crosses has been reported by several authors (e.g., Santalla et al. 2005; Welsh et al. 1995).

Our results suggest that Phaseolin and other seed proteins are linked to favorable alleles at loci affecting seed size. These markers have good potential for improving the results of the selection and should be considered by breeders as a strategy for germplasm enhancement and to avoid the reduced performance of the inter-gene pool populations by means of a marker-assisted selection to identify Andean alleles in the progenies that posses a positive effect on seed size.

The availability of tags for genes that control seed size can be useful in studies of bean evolution. In addition to phaseolin type, seed size is one of the major differentiating characters between Mesoamerican and Andean *Phaseolus vulgaris* beans (Gepts 1998). The availability of genetic markers for genes involved in seed size control in beans will certainly aid physiological and evolutionary studies centered on this important agronomic character.

There are several implications of our results to breeding programs. Based on molecular markers, crosses involving Mesoamerican × Andean genotypes were found to be the most divergent and may generate more significant progress from selection. The transfer of polygenic traits such as seed size potential of Andean types to Mesoamerican types may, however, be a problem, because distorted segregations make it more difficult to recover certain recombinants. This may explain in part why it has been difficult to transfer quantitative traits between cultivars of Mesoamerican and Andean origin.

We conclude that selection for seed size may lead to molecular changes in protein frequencies in the selected populations. Results from this work are relevant to ongoing common bean breeding programs that are aimed at developing beans with a good seed commercial quality, in which protein markers can be used as tools to select.

Acknowledgments Research was supported by the projects AGF97-0324 and AGL2005-01268/AGR from the Spanish Government, PGIDIT02RAG40301PR from the Galician Government (Spain), and EU-FEDER Funds. A. M. González thanks her fellowship to Ministry of Education, Culture and Sports of Spain. M. De la Fuente is grateful to the Xunta de Galicia for awarding her a fellowship grant. The authors thank also to Diputación Provincial de Pontevedra (Spain) for farm facilities.

References

Blair MW, Iriarte G, Beebe S (2006) QTL analysis of yield traits in an advanced backcross population derived from a

cultivated Andean × wild common bean (*Phaseolus vulgaris* L.) cross. Theor Appl Genet 112:1149–1163. doi:10.1007/s00122-006-0217-2

Brown JWS, Ma Y, Bliss FA, Hall TC (1981) Genetic variation in the subunits of globulin-1 storage protein of French bean. Theor Appl Genet 59:83–88. doi:10.1007/BF00285895

Brown JWS, Osborn TC, Bliss FA, Hall TC (1982) Bean lectins. Theor Appl Genet 62:361–367. doi:10.1007/BF00275105

Butrón A, Tarrío R, Revilla P, Ordás A, Malvar RA (2005) Molecular changes in the maize composite EPS12 during selection for resistance to pink stem borer. Theor Appl Genet 6:1044–1051. doi:10.1007/s00122-005-1923-x

Damerval C, Maurice A, Josse JM, de Vienne D (1994) Quantitative trait loci underlying gene product variation: a novel perspective for analyzing regulation of gene expression. Genetics 137:289–301

De Koeyer DL, Phillips RL, Stuthman DD (2001) Allelic shifts and quantitative trait loci in a recurrent selection population of oat. Crop Sci 41:1228–1234

Delaney DE, Bliss FA (1991) Selection for increased percentage phaseolin in common bean. 1. Comparison of selection for seed protein alleles and S1 family recurrent selection. Theor Appl Genet 81:301–305. doi:10.1007/BF00228668

Devicente MC, Tanksley SD (1993) QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics 134:585–596

Diaby M, Casler MD (2005) RAPD marker variation among divergent selections for neutral detergent fiber concentration in four smooth bromegrass populations. Crop Sci 45:27–35

Eagen KA, Goldman IL (1996) Assessment of RAPD marker frequencies over cycles of recurrent selection for pigment concentration and percent solids in red beet (*Beta vulgaris* L.). Mol Breed 2:107–115

Edwards MC, Stuber CW, Wendel JF (1987) Molecular marker-facilitated investigations of quantitative-trait loci in maize. I. Numbers, genomic distribution, and types of gene action. Genetics 116:113–125

Freyre R, Skroch P, Geffroy V, Adam-Blondon AF, Shirmohamadali A, Johnson W, Llaca V, Nodari R, Pereira P, Tsai SM, Tohme J, Dron M, Nienhuis J, Vallejos CE, Gepts P (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. doi:10.1007/s001220050964

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

Gepts P, Osborn T, Rashka K, Bliss F (1986) Phaseolin protein variability in wild forms and landraces of the common bean (*Phaseolus vulgaris* L.): evidence for multiple centers of domestication. Econ Bot 40(4):451–468

González AM, Rodiño AP, Santalla M, De Ron AM (2009) Genetics of intra-gene pool and inter-gene pool hybridization for seed traits in common bean (*Phaseolus vulgaris* L.) germplasm from Europe. Field Crops Res 112:66–76. doi:10.1016/j.fcr.2009.02.003

Guzman PS, Lamkey KR (2000) Effective population size and genetic variability in the BS11 maize population. Crop Sci 40:338–346

Guzman-Maldonado SH, Martínez O, Acosta-Gallegos JA, Guevara-Lara F, Paredes-López O (2003) Putative quantitative trait loci for physical and chemical components of common bean. Crop Sci 43:1029–1035

Hartana A (1983) Genetic variability on seed protein levels associated with two phaseolin types in common bean (*Phaseolus vulgaris* L.). M.S. dissertation, University of Wisconsin

Heredia-Diaz O, Alsirt A, Darrah LL, Coe EH (1996) Allelic frequency changes in the MOSCSSS maize synthetic in response to bi-directional selection for rind penetrometer resistance. Maydica 41:65–67

Johnson WC, Gepts P (2002) The role of epistasis in controlling seed yield and other agronomic traits in an Andean × Mesoamerican cross of common bean (*Phaseolus vulgaris* L.). Euphytica 125:69–79. doi:10.1023/A:1015775822132

Johnson WC, Menéndez C, Nodari R, Koinange EMK, Magnusson S, Singh SP, Gepts P (1996) Association of a seed weight factor with the phaseolin seed storage protein locus across genotypes, environments, and genomes in *Phaseolus-Vigna* spp.: Sax (1923) revisited. J Agric Genomics 2 (previously J Quant Trait Loci). http://www.cabi-publishing.org/gateways/jag/papers96/paper596/indexp596.html

Kahler AL (1983) Effect of half-sib and S1 recurrent selection for increased grain yield on allozyme polymorphisms in maize. Crop Sci 23:572–576

Kaplan NL, Hudson RR, Langley CH (1989) The "hitchhiking effect" revisited. Genetics 123:887–899

Koenig R, Gepts P (1989a) Allozyme diversity in wild *Phaseolus vulgaris*: further evidence for two major centers of diversity. Theor Appl Genet 78:809–817. doi:10.1007/BF00266663

Koenig R, Gepts P (1989b) Segregation and linkage of genes for seed proteins, isozymes, and morphological traits in common bean (*Phaseolus vulgaris* L.). J Hered 80:455–459

Labate JA, Lamkey KR, Lee M, Woodman WL (1999) Temporal changes in allele frequencies in two reciprocally selected maize populations. Theor Appl Genet 99:1166–1178. doi:10.1007/s001220051321

Lioi L (1991) Electrophoretic variation and geographical distribution of the seed protein phytohemagglutinin in cultivated *Phaseolus vulgaris* L. J Genet Breed 45:97–102

Mennella G, Sanajà VO, D'Alessandro A, Milone M, Perrone D (2003) HPLC analyses of seed storage proteins reveal polymorphism in Italian common bean (*Phaseolus vulgaris* L.) ecotypes. Euphytica 134:85–95. doi:10.1023/A:1026127004224

Miles C, Wayne M (2008) Quantitative trait locus (QTL) analysis. Nat Educ 1:1

Motto M, Soressi GP, Salamini F (1978) Seed size inheritance in a cross between wild and cultivated common beans (*Phaseolus vulgaris* L.). Genetica 49:31–36. doi:10.1007/BF00187811

Murray JD, Michaels TE, Cardona C, Schaafsma AW, Pauls KP (2004) Quantitative trait loci for leafhopper (*Empoasca fabae* and *Empoasca kraemeri*) resistance and seed weight in the common bean. Plant Breed 123(6):474–479. doi:10.1111/j.1439-0523.2004.01020.x

Mutlu N, Miklas P, Reiser J, Coyne D (2005) Backcross breeding for improved resistance to common bacterial blight in pinto bean (*Phaseolus vulgaris* L.). Plant Breed 124:282–287. doi:10.1111/j.1439-0523.2005.01078.x

Nienhuis J, Singh SP (1988) Genetics of seed yield and its components in common bean (*Phaseolus vulgaris* L.) of Middle-American origin. I. General combining ability. Plant Breed 101:143–154. doi:10.1111/j.1439-0523.1988.tb00280.x

Nodari R, Tsai SM, Gilbertsont RL, Gepts P (1993) Towards an integrated linkage map of common bean. II. Development of an RFLP-based linkage map. Theor Appl Genet 85:513–520. doi:10.1007/BF00220907

Osborn TC, Burow M, Bliss HA (1988) Purification and characterization of arcelin seed protein from common bean. Plant Physiol 86:399–405

Park SO, Coyne DP, Jung G, Skroch PW, Arnaud SE, Steadman JR, Ariyarathne HM, Nienhuis J (2000) Mapping of QTL for seed size and shape traits in common bean. J Am Soc Hortic Sci 125:466–475

Paterson AH, Lander ES, Had JD et al (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335:721–726. doi:10.1038/335721a0

Pollak LM, Gardner CO, Parkhurst AM (1984) Relationships between enzyme marker loci and morphological traits in two mass-selected maize populations. Crop Sci 24(6): 1174–1179

Pustzai A, Clarke EMW, King TP, Stewart JC (1979) Nutritional evaluation of kidney beans (*Phaseolus vulgaris*): chemical composition, lectin content, and nutritional value of selected cultivars. J Sci Food Agric 30:843–848

Roach DA, Wulff RD (1987) Maternal effects in plants. Ann Rev Ecol Syst 18:209–215

Romero Andreas J, Yandell BS, Bliss FA (1986) Bean arcelin I. Inheritance of a novel seed protein of *Phaseolus vulgaris* L. and its effect on seed composition. Theor Appl Genet 72:123–12810. doi:1007/BF00261467

Santalla M, Rodiño AP, González AM, Monteagudo AB, De Ron AM (2005) Improvement of large-seeded common bean cultivars under sustainable cropping system in Spain. Euphytica 142:85–95. doi:10.1007/s10681-005-0816-z

SAS (2002) The SAS System. SAS online Doc-HTML Format-Version nine. Institute Inc, Cary

Sax K (1923) The association of size differences with seed coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics 8:552–560

Schaffer HE, Yardley D, Anderson WW (1977) Drift or selection: a statistical test of gene frequency variation over generations. Genetics 87:371–379

Singh SP (1992) Common bean improvement in the Tropics. Plant Breed Rev 10:199–269

Singh SP, Urrea CA (1994) Selection for seed yield and other traits among early generations of inter- and intraracial populations of the common bean. Rev Bras Genet 17:299–303

Singh SP, Gepts P, Debouck DG (1991a) Races of common bean (*Phaseolus vulgaris*, Fabaceae). Econ Bot 45:379–396

Singh SP, Nodari RO, Gepts P (1991b) Genetic diversity in cultivated common bean I. Allozymes. Crop Sci 31:19–23

Singh SP, Teran H, Muñoz CG, Osorno JM (2002) Selection for seed yield in Andean intra-gene pool and Andean × Middle American inter-gene pool populations of common bean. Euphytica 127:437–444. doi:10.1023/A:1020317608553

Stephan W, Wiehe THE, Lenz MW (1992) The effect of strongly selected substitutions on neutral polymorphism: analytical results based on diffusion theory. Theor Popul Biol 41:237–254

Stuber CW, Moll RH, Goodman MM, Schaffer HE, Weir BS (1980) Allozyme frequency changes associated with selection for increased grain yield in maize (*Zea mays* L.). Genetics 95:225–236

Tanksley SD (1993) Mapping polygenes. Annu Rev Genet 27:205–233

Vallejos CE, Chase CD (1991a) Linkage between isozyme markers and a locus affecting seed size in *Phaseolus vulgaris* L. Theor Appl Genet 81:413–419. doi:10.1007/BF00228685

Vallejos CE, Chase CD (1991b) Extended map for the phaseolin linkage group of *Phaseolus vulgaris* L. Theor Appl Genet 82:353–357. doi:10.1007/BF02190622

Weeden NF, Liang CY (1985) Detection of a linkage between white flower color and EST-2 in common bean. Annu Rep Bean Improv Coop 28:87–88

Welsh W, Bushuk W, Roca W, Singh SP (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. doi:10.1007/BF00220874

Wilson SR (1980) Analyzing gene frequency data when the effective population size is finite. Genetics 95:489–502