Field Crops Research 112 (2009) 66–76

Contents lists available at ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

Genetics of intra-gene pool and inter-gene pool hybridization for seed traits in common bean (*Phaseolus vulgaris* L.) germplasm from Europe

A.M. González *, A.P. Rodiño, M. Santalla, A.M. De Ron

Plant Genetic Resources Department, Misión Biológica de Galicia, CSIC, P.O. Box 28, 36080 Pontevedra, Spain

ARTICLE INFO

Article history:
Received 14 November 2008
Received in revised form 3 February 2009
Accepted 9 February 2009

Keywords:
Phaseolus vulgaris
Large-seeded alleles
Combining ability
Segregation distortion
Phs locus

ABSTRACT

Due to adaptation to new ecological and manmade conditions, the large diversity evolved in the European common bean germplasm is of particular interest for plant breeding. The knowledge of the genetic relationships within and among races and gene pools and their performance per se will provide bean breeders with a starting point in designing crosses using contrasting and complementary parents to broaden the genetic base within the different commercial classes. A genetic study of seed size variation and protein markers in progeny derived from 16 intraracial, interracial and inter-gene pools European common bean populations was conducted. General and specific combining ability (GCA and SCA) values were significant for seed weight, indicating that both additive and nonadditive genetic effects were involved in conditioning seed weight. Interracial populations showed transgressive values due to the accumulation of large-seeded alleles. Genetic variation inside Andean germplasm, and Chile and Peru races in particular, exhibited useful genetic progress in these populations, providing lines with a large seed size, and so, an excellent market potential. The distribution of incompatibility between both gene pools (Mesoamerican and Andean) of the common bean was explored. Inter-gene pool populations provided lower means of inbred segregants than the mid-parent value. Therefore, a good option it would be select for large seed size according to a recurrent or congruity inbred-backcrossing selection programs. Analysis of allele markers frequencies in inter-gene populations showed segregation distortion with a higher than expected frequency of alleles from the Mesoamerican gene pool, many of which were fixed in the F_7 lines. The presence of a great percentage of markers that showed segregation distortion in these populations (87%) indicated that this phenomenon can be amplified by using distance related common bean genotypes. In addition, a high percentage of heterozygotes for the *Phs* locus (for the seed storage protein phaseolin) was found, which suggest that the Andean homozygous TT could not be expressed in Mesoamerican genetic background due to the action of some form of female specific mechanisms that affected gene exchange between parental germplasm in inter-gene pool populations. The present work provides useful information in the establishment of large seed size germplasm that could have a great deal of interest among breeders and may offer some possibilities to exploit existing variation within and between common bean races.

1. Introduction

Cultivated common bean (*Phaseolus vulgaris* L.) was domesticated independently within two centers of diversity giving rise to two gene pools: Mesoamerican and Andean (Gepts et al., 1986; Koenig and Gepts, 1989a; Singh et al., 1991a,b,c; Beebe et al., 2000). The differences between Mesoamerican and Andean gene pools of common bean include plant morphology, seed size, phaseolin (seed storage protein) patterns, isozymes and DNA polymorphisms. The study of a wide range of molecular-genetic markers has confirmed that high levels of polymorphism are present in Mesoamerican × Andean mapping populations. Substantial levels of polymorphism have also been found within both gene pools (e.g. 60–70% of probes assayed, compared with 80–90% in inter-gene pool crosses) (Gepts et al., 1993). Singh and Gutièrrez (1984) considered both gene pools to be well on the way to speciation and postulated the existence of genes for incompatibility between them, whereas Gepts (1988a,b) considered that the marked differences between both pools were the result of their relative geographical isolation. The availability of genetic variation from inter-gene pool hybridizations offers a potential opportunity to breeders of common bean. It may be possible to recombine the desirable traits of each gene pool, such as the greater yield of Mesoamerican genotypes with the larger seed size of Andean

* Corresponding author. Tel.: +34 986854800; fax: +34 986841362.
E-mail address: amgonzalez@mbg.cesga.es (A.M. González).

0378-4290/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.fcr.2009.02.003

genotypes, into improved cultivars. However, many other features are also inherited, along with the desirable ones. A breeding program lasting for many years may be necessary to eliminate the unwanted characteristics without losing the desired one(s). Several crosses between Mesoamerican and Andean gene pools of common bean have been studied, but those have generally not involved systematic combinations of a wide range of genotypes of each gene pool.

Bean gene pools have further been divided into races, where the term 'race' is used to denote a group of related landraces (Gepts, 1988b). There are three Mesoamerican (Mesoamerica, Durango, and Jalisco) and three Andean races (Peru, Nueva Granada, and Chile). Members of each race have distinct and specific physiological, agronomic, biochemical and molecular characteristics and differ from other races in the allelic frequencies at specific loci (Singh et al., 1991a,b,c). Because each race may have been domesticated from different wild populations in ecologically and geographically discrete regions under discrete selection pressures, different mechanisms and genes for expression of traits could exist for each race.

Gepts (1988b) suggested that the diversity of the European common bean was reduced due to a strong founder effect during its introduction in Europe. However, recent studies show that the reduction of genetic diversity was not as great as expected and gene flow between both pools has been suggested (Masi, 2001; Santalla et al., 2002; Papa et al., 2005). The level of genetic variation has not been eroded since the introduction of the common bean from the American centers of domestication to the Iberian Peninsula (Spain and Portugal). Instead, obvious signs of introgression between both gene pools were observed, mainly among white-seeded genotypes. Large diversity of morphological traits has been observed in the Iberian Peninsula (De la Cuadra et al., 2000). The Iberian Peninsula germplasm is therefore more complex than previously thought, and contains additional diversity that remains to be explored for genetic and breeding proposes (Rodiño et al., 2006).

Consumers have progressively shown specific preferences for various combinations of size and shape of bean seeds, and the market reflects this trend by giving preference to types of good quality rather than high yield. In Europe, white large-seeded varieties have commercial advantage over existing varieties. Some Spanish dry bean varieties have acquired considerable importance due to their good seed quality and these varieties are protected by law. Selection to maintain or increase seed size could have a great deal of interest among breeders and may offer some possibilities of expanding commercial production and consumption.

Strategies employed by dry bean breeders to improve yield include early generation testing (EGT) and selection based on yield components (Nienhuis and Singh, 1988). Seed size (weight) is one of the major yield components in dry bean breeding programs. However, several authors (e.g. Nienhuis and Singh, 1988; Singh, 1995) reported a significant negative correlation between seed size and seed yield. Mesoamerican germplasm possess smaller seed size (100-seed weight $\leq$25 g) than its Andean counterpart (100-seed weight $\geq$25 g) (Gepts et al., 1986). A significant genetic gain could require the development of populations from crosses between both gene pools, but phenotypic abnormalities, (Gepts and Bliss, 1985; Koinange and Gepts, 1992), and the appearance of undesirable segregants in later generations, including diminished seed yield potential, may occur. Gene exchange between parents from the same race should proceeds without metabolic or physiological hindrance because of evolutionary similarity, but it generally results in reduced genetic variation and few polymorphic markers (Welsh et al., 1995; Gepts, 1998).

Several markers have been used to study the genetics of seed weight in common bean, including seed pigmentation (*P* gene) (Sax, 1923), isozymes (Vallejos and Chase, 1991; Vallejos et al., 1992), seed protein markers (Hartana, 1983; Delaney and Bliss, 1991; Johnson et al., 1996), restriction fragment length polymorphisms (RFLPs) (Nodari et al., 1992; Park et al., 2000) and random amplified polymorphic DNA (RAPD) markers (Haley et al., 1994). The seed-storage protein, phaseolin (Pha) has been proved to be an excellent evolutive marker in common bean. Accessions from Mesoamerica have the S and B phaseolin patterns (with some wild accessions exhibiting the M phaseolin pattern), while accessions from the Andes have primarily the T phaseolin pattern, with some accessions exhibiting C, and H, and wild common beans have A, J and I patterns (Gepts and Bliss, 1986; Gepts et al., 1986). Phaseolin protein constitutes about 40% of the major globulin storage protein of common bean seeds (Ma and Bliss, 1978). The globuline-2 (G2) protein fraction constitutes between 5% and 12% of the total seed protein, and some of these proteins have been suggested to be lectins. Studies of the seed storage proteins of many plants have suggested that the numerous storage protein subunits, distinguishable by molecular weight and isoelectric point, are the products of different but evolutionary related genes clustered at a few genomic sites. Phaseolin proteins are made up of three different subunits with molecular weights that vary from 43 kDa to 54 kDa (Osborn et al., 1988). The *Phs* locus on B7 linkage group (Freyre et al., 1998) codes for phaseolin protein, which harbours an estimated six to nine sequences (Talbot et al., 1984), and is linked to genes encoding other seed storage proteins.

Isozymes markers have been used also as a tool to differentiate between genotypes belonging to the Mesoamerican and Andean gene pools (Singh et al., 1991c; Debouck et al., 1993; Santalla et al., 2002) and for estimating the genetic relationships among market classes (Bassiri and Adams, 1978). Weeden (1984, 1986) and Weeden and Liang (1985) determined that the polymorphisms for *Ribulose biphosphate carboxylase*, *Shikimate dehydrogenase*, *Peroxidase*, *Malic*, *Glucose phosphate isomerase*, N-*Acetyl Glucosaminidase* and *Adenylate Kinase* were controlled by single genes. Sprecher (1988) determined that *diaphorase* polymorphism was controlled by two tightly linked genes, *Diap-1* and *Diap-2*, and Koenig and Gepts (1989b) identified two additional polymorphisms for *Leucine aminopeptidase* and *Malate dehydrogenase*, coded by *Lap-3* and *Mdh-1*, respectively. *Diap-1* and *Diap-2* loci and *Rbc* and *Me* loci ($r < 30$ cM) are unlinked and therefore represent different regions of the genome.

The success of a breeding program depends crucially on the choice of parents to be involved in hybridization. Breeders generally utilize parents that make possible to get populations with a high average associated with great variability for the traits under selection. General and specific combining abilities (GCA and SCA) are parameters employed to choose promising segregating populations. These parameters allow identifying populations which are potentially more useful to release variability in the segregating generations. Spanish landraces deserve attention since they could be the greatest source of variation for the genetic improvement of large and extra-large white-seeded bean varieties worldwide. There is a large base of germplasm that can be utilized by breeders in the production of promising lines. In this paper, we explore a wide set of crosses within and among races to identify germplasm sources useful for enhancing seed size in a breeding program. The objectives of this work were to: (1) propose the best genotype combinations to allow selection of potential parents or populations with improved seed characteristics and (2) analyze the genotypic performance and the progeny segregation of crosses within and between the two major common bean gene pools.

2. Materials and methods

2.1. Genetic material

Thirty dry bean inbred lines with allelic differences at six enzyme loci, phaseolin and other seed proteins, and differences in

Table 1
Characteristics of common bean lines used to obtain the segregating populations.

Identification	Race combination[a]	Market class combination[b]	[c]Phaseolin combination[a]	Size combination[d]	Growth habit combination[e]
Intraracial populations					
PHA-20-07/PHA-159-11	M × M	SW × SW	S × B	Medium × Small	III × II
PHA-267-20/PHA-257-23	NG × NG	WK × WK	T × T	Large × Large	I × I
PHA-338-19/PHA-306-21	P × P	DG × DG	H × H	Large × Large	I × I
Interracial populations					
PHA-452-01/PHA-119-01	NG × P	F × DG	T × H	Large × Medium	IV × II
PHA-257-01/PHA-323-02	NG × CH	WK × BG	T × C	Large × Large	I × IV
PHA-257-06/PHA-306-01	NG × P	WK × DG	T × H	Large × Large	I × I
PHA-257-10/PHA-306-11	NG × P	WK × DG	H × T	Large × Large	I × I
PHA-267-18/PHA-338-27	NG × P	WK × DG	T × H	Large × Large	I × I
PHA-272-01/PHA-257-01	P × NG	PC × WK	H × T	Medium × Large	IV × I
PHA-272-02/PHA-257-04	P × NG	PC × WK	H × T	Medium × Large	IV × I
Inter-gene pool populations					
PHA-159-09/PHA-257-08	M × NG	SW × C	S × T	Small × Large	II × I
PHA-159-12/PHA-267-15	M × NG	SW × WK	B × T	Small × Large	II × I
PHA-159-13/PHA-267-18	M × NG	SW × WK	S × T	Small × Large	II × I
PHA-159-14/PHA-257-14	M × NG	SW × WK	S × T	Small × Large	II × I
PHA-159-10/PHA-269-12	M × NG	SW × DRK	S × T	Small × Large	II × I
PHA-159-08/PHA-269-13	M × NG	SW × DRK	S × T	Small × Large	II × I

[a] Mesoamerican races—M: Mesoamerica; Andean races—NG: Nueva Granada; P: Peru; CH: Chile.
[b] Market classes—SW: Small White; WK: White Kidney; DG: Dark Garbanzo; F: Favada; BG: Brown Garbanzo; PC: Purple Caparron; C: Canellini; DRK: Dark Red Kidney.
[c] Mesoamerican phaseolin—S: Sanilac, B: Boyaca; Andean phaseolin—T: Tendergreen; C: Contender; H: Huanchaco.
[d] Large (weight of 100 seeds less than 25 g); Medium (weight of 100 seeds between 25 and 40 g); Small (weight of 100 seeds larger than 40 g).
[e] I: determinate, erect; II: indeterminate, erect; III: indeterminate, non-climber or semi-climbing, prostrate, and IV: indeterminate, climbing.

seed size were selected from landraces of the Misión Biológica de Galicia-National Spanish Research Council (MBG-CSIC) breeding collection. These landraces were collected in the main dry bean production regions of the Iberian Peninsula. Several morphological and agronomic traits, phaseolin protein, and allozymes were used to group the inbred lines in different market classes according to Voysest and Dessert (1991), the Andean and Middle American gene pools, and their affiliations with the races described by Singh et al. (1991a) in the Americas (Table 1). Large-seeded bean genotypes belonging to the races Nueva Granada (nine White Kidney, one Canellini, two Dark Red Kidney and one Favada genotypes), Chile (Brown Garbanzo) and Peru (two Purple Caparron, and six Dark garbanzo genotypes) from the Andean gene pool, and the small-seeded genotypes belonging to the race Mesoamerica (eight Small White genotypes) from the Mesoamerican gene pool were included to perform intra- (three), interracial (seven) and inter-gene (six) biparental crosses.

F_1 seeds, together with their respective parents, were grown in the field to ascertain hybridity origin and produce F_2 seed. The F_2 seeds from F_1 plans of 16 cross combinations were grown and formed 16 F_2 segregation populations. These sixteen crosses were advanced to the field nursery for two more generations until F_4 generation. That is, for each cross after F_2, a random sample of seeds was taken to obtain the next generation. Most of the progeny grew normally; only few unhealthy plants were not viable. This was followed by single plant harvests selected for seed commercial quality (proportion of seed coat, seed size and quality seed aspect) combined with seed yield from F_4 to F_5. They were grown in plant-to-progeny rows in the F_6 and F_7. All plants within each plot were bulk-harvested. From F_2 to F_7 seed was stored at 4–5 °C temperature and 50% humidity.

Measurements and protein analyses have been made in each seed and in all generations.

2.2. Seed size and morphological data

Four quantitative seed traits were scored for each individual seed: weight (g seed^{-1}), length (mm), defined as the longest distance across the seed parallel to the hilum, height (mm), as the longest distance perpendicular to length, and width (mm), measured as the longest distance across the hilum seed. The seed coat color was determined by using a binomial scale, in which 0 = white seed coat, and 1 = colored seed coat. The 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.

2.3. Seed storage protein analysis

For storage protein analysis, a portion of the seed was manually removed, prior to germination, and ground into a fine dust. The flour sample was suspended at room temperature for at least 30 min 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 microliter samples 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) polyacrilamide slab gels (height, 16 cm; width, 18 cm) and carried out 25 mA at loading (approximately 1 h) and thereafter to 30 mA until the separation was completed. Proteins were stained with Coomassie brilliant blue R-250.

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), b4 and b5 polypeptides (58–55 kDa), b6 and b7 polypeptides (42–41 kDa), and b11, b12 and b13 polypeptides or lectins (33–31 kDa).

2.4. Isozyme analysis

The other seed portion was sown for isozyme analysis. A combination of starch gel electrophoresis and enzyme activity staining was used to detect polymorphisms 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.

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, 1989b). 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 $Mdh\text{-}2^{100}$; while Dark Red Kidney exhibits the following genotype: $Skdh^{100}$, Me^{98}, Rbc^{98}, Prx^{98}, $Diap\text{-}1^{100}$, $Diap\text{-}2^{105}$, $Mdh\text{-}1^{103}$ and $Mdh\text{-}2^{102}$. The isozyme and phaseolin characters are embryonic characteristics, so reflect the F_n plant genotype.

2.5. Statistical analysis

Statistical analyses were carried out using the Statistical Analysis System (SAS v9: SAS Institute Inc., Cary, NC, USA, 2002). Simple correlation coefficients were calculated among seed size traits using Spearman rank correlations and the associated significance levels were calculated using the PROC CORR SPEARMAN procedure in SAS. Mann–Whitney *U* and Kruskal–Wallis non-parametric tests were performed to test the hypothesis that the medians of the progeny were equal to the parents and among them in each one of the seed quantitative traits studied. Mean values of traits were compared using LSD at the 0.05 level of probability.

Segregation data for seed color pigmentation, protein and isozyme marker classes were tested for goodness-of-fit to a monogenic segregation (3:1 or 1:2:1) model using the χ^2 statistic test ($P \geq 0.05$) and according to the generation studied. χ^2 statistic was used to test the goodness-of-fit of the observed ratio to expected ratios before the selection and reflect the impact of the selection that was performed from F_5 generation.

The GCA (g_i) was calculated as the average of all progeny having this particular race as one parent, the value being expressed as a deviation from the overall mean of crosses. Estimates g_i close to zero indicates that the race does not differ from the general mean of all crosses. On the other hand, positive or negative g_i values indicate superior or inferior races compared with the other races employed.

The GCA of each race was calculated as follows:

$$g_i = \left[\frac{T_i}{(n-2)}\right] - \left[\sum \frac{T}{n}(n-2)\right],$$

where *i* represents a specific race, $\sum T$ represents the overall mean of crosses, and *n* the number of races.

Each cross has an expected value (the sum of GCAs of its two parental races). However, each cross may deviate from the expected value to a greater or lesser extent, the deviation being the specific combining ability (SCA) (s_{ij}) of the two races in combination. Low s_{ij} values indicate that the respective races exhibited the seed size predicted by their parent's general combining abilities. However, positive or negative s_{ij} values indicate that the seed size of a particular cross is respectively better or worse than expected, based on the parental GCA. The mathematical representation of this relationship for each cross is

$$X_{ij} = \bar{X} + g_i + g_j + s_{ij},$$

where $\bar{X}$ is the general mean and g_i and g_j are the general combining ability estimates of the parents, and s_{ij} is the statistically unaccounted for residual or specific combining ability.

Significant differences among the GCA and SCA effects were tested using *F*-values.

3. Results

3.1. Correlation coefficients among seed size traits

All crosses showed significant phenotypic correlations (<0.01) among all phenotypic traits studied, except for population PHA-159-12/PHA-267-15, which showed a non-significant correlation between seed length and width, and height and width. In all crosses, independently of origin of the parents, the lowest correlations values were between length–width and height–width (data not shown).

3.2. Genetic changes and combining ability in the segregating generations

Estimates of general combining ability (GCA) of each race for seed size characters are presented in Table 2. The GCA effects for seed weight were significant for all races, one of which, Mesoamerica race, was negative. Large positive GCA values indicated effective transmission of genes from parents to their offspring for these traits. The positive values presented by the races Chile and Peru indicated that they are good candidates to improve large seed size. It was also observed that the highest means were presented by populations of which both parents showed positive GCA. Table 2 also presents estimates of the specific combining ability (s_{ij}) of the six race combinations. The SCA effects for seed weight, length and seed height were significant for four of the six combinations. The SCA within interracial crosses Peru × Nueva Granada ranged from 0.06 to 0.17 for seed weight, indicating that is possible to find differences between parents, that is, some parents complement each other better than others. Intraracial crosses Mesoamerica × Mesoamerica and Peru × Peru might not combine well among themselves to improve seed size, as indicated by their zero values. When the SCA is not significant, the performance of the cross could be adequately predicted on the basis of GCA.

Intraracial populations showed a narrow range of variation for all seed traits studied, except in the population PHA-267-20/PHA-257-23 (Table 3). Genetic variability within bean races for all characters is limited. Thus, it is evident that the use of only one source of variability in intraracial populations would not be adequate for developing lines with large seed. In contrast, interracial populations reflected a substantial variability and constituted a source of useful genetic variability to maximize the genetic gain of selection. However, inter-gene pool populations displayed surprisingly lower overall genetic variability compared to interracial populations. The observed variability was lower than expected. Therefore, other breeding strategies are needed to increase the genetic variability in these crosses and sort out the factors that have discouraged inter-gene pool recombination.

The medians for seed weight, length, height and width of the segregating generations were significantly different from mid-parent value (some relevant populations are presented in Table 4). Outliers were observed in all populations for seed trait medians, suggesting that both parents carry alleles for large and small seed size. The interracial populations have given a better outcome than the intraracial populations for the seed traits studied, exhibiting greater potential to increase the seed size. The crosses White Kidney × Brown Garbanzo (PHA-257-01/PHA-323-02) and White Kidney × Dark Garbanzo (PHA-257-06/PHA-306-01 and PHA-267-

Table 2
Mean values for $F_{3:7}$ generations and parents, estimates of the general combining ability (g_i) and specific combining ability effects (s_{ij}) for seed weight, length, height and width.

Races/cross	N^a	N^b	Seed weight (g seed^{-1})			Length (mm)		Height (mm)		Width (mm)	
			Parents	N^c	$F_{3:7}$	Parents	$F_{3:7}$	Parents	$F_{3:7}$	Parents	$F_{3:7}$
Mean											
Mesoamerica	7	40	0.27a	3884	0.34a	10.88a	11.58a	7.21a	7.50a	5.28a	5.54a
Nueva Granada	14	75	0.58b	7969	0.54b	16.11c	14.30b	7.86a	7.99a	6.34b	6.10b
Peru	7	40	0.51b	3997	0.78b	13.35b	14.63b	8.04b	8.28b	6.59b	6.48b
Chile	1	5	0.70b	448	0.68c	15.32b	16.35c	8.80b	8.85b	7.64c	7.21c
Total mean			0.50			13.83		7.94		6.13	
g_i											
Mesoamerica	7	3924[d]	−0.15**			−2.08**		−0.44**		−0.54**	
Nueva Granada	14	8044	0.06*			0.24		0.08		−0.03	
Peru	7	4037	0.11**			1.40**		0.37*		0.36*	
Chile	1	453	0.21**			3.00**		0.71**		0.96**	

Races/cross	N^a	N^b	Seed weight (g seed^{-1})			Length (mm)		Height (mm)		Width (mm)	
			Parents	Ind.	$F_{3:7}$	Parents	$F_{3:7}$	Parents	$F_{3:7}$	Parents	$F_{3:7}$
Mean											
M × NG	6	60	0.42b	2916	0.36b	13.65	11.96b	7.55a	7.64a	5.25a	5.57a
P × NG	6	60	0.59c	3303	0.66d	14.53	15.20d	7.74a	8.38c	6.15b	6.40b
NG × CH	1	10	0.62d	448	0.78e	15.77	16.76e	8.54b	8.75d	7.38c	7.33d
M × M	1	10	0.33a	968	0.30a	11.09	11.07a	7.70a	7.49a	5.53a	5.88a
P × P	1	10	0.60c	694	0.56c	14.20	14.02c	8.57b	8.31c	7.13c	6.97c
NG × NG	1	10	0.62d	1302	0.61c	15.70	15.58d	7.76a	8.05b	6.42b	6.45b
s_{ij} (*range of crosses*)											
M × NG	6	2976[e]		−0.06* (−0.1, −0.006)		−0.69** (−1.7, 0.12)		−0.06 (−0.4, 0.3)		−0.07 (−0.7, 0.11)	
P × NG	6	3363		0.11** (0.06, 0.17)		1.16** (3.26, 0.81)		0.44** (0.72, −0.08)		0.32* (0.9, −0.16)	
NG × CH	1	458		0.18**		1.93**		0.58**		0.68**	
M × M	1	978		0		−0.27		0.18		0.43*	
P × P	1	704		0		−0.30		0.21		0.67**	
NG × NG	1	1312		0.11**		1.62**		0.37*		0.57**	

Means followed by the same letter within a column are not significantly at $P = 0.05$ level.
[a] Number of crosses involved.
[b] Number of parental involved in mean estimation.
[c] Number of progeny involved in mean estimation.
[d] Number of progeny involved in g_i estimation.
[e] Number of progeny involved in s_{ij} estimation.
* Significance at $P = 0.05$ level.
** Significance at $P = 0.01$ level.

Table 3
Mean performance of all generations, standard deviation (S), and range of variation for seed traits in the 16 intraracial, interracial and inter-gene pool common bean populations studied.

Populations	N^a	Weight			Length			Width			Height		
		Mean	S	Range	Mean	S	Range	Mean	S	Range	Mean	S	Range
Intraracial populations													
PHA-20-07/159-11	968	0.32	0.059	0.14–0.52	10.9	0.78	8.5–13.3	7.4	0.54	5.5–9.4	5.7	0.66	3.8–7.6
PHA-267-20/257-23	1302	0.62	0.100	0.33–1.08	15.1	1.25	9.2–19.3	7.8	0.79	4.7–11.6	6.1	0.70	3.4–8.1
PHA-338-19/306-21	694	0.53	0.073	0.48–0.98	14.0	1.23	10.2–19.1	8.3	0.76	5.9–10.4	6.9	0.71	3.2–9.2
Interracial populations													
PHA-452-01/119-01	220	0.70	0.190	0.36–1.41	17.3	2.23	11.7–23.5	8.4	0.99	5.4–10.7	6.5	0.91	4.2–9.6
PHA-257-01/323-02	448	0.74	0.196	0.27–1.28	16.8	1.42	11.9–20.1	8.8	1.02	5.8–11.1	7.3	0.83	4.3–9.0
PHA-257-06/306-01	1323	0.71	0.175	0.26–1.32	16.9	1.29	12.0–21.5	8.5	0.73	5.5–11.9	6.9	0.70	4.7–10.1
PHA-257-10/306-11	824	0.53	0.150	0.17–1.07	14.9	1.60	9.4–20.4	7.9	0.87	4.7–10.5	6.2	0.82	3.7–8.3
PHA-267-18/338-27	627	0.66	0.152	0.15–1.28	16.3	1.85	10.4–20.8	8.5	0.90	5.3–11.1	7.0	0.84	4.9–10.6
PHA-272-01/257-01	155	0.60	0.127	0.34–1.02	14.9	1.23	11.3–18.2	8.2	0.72	6.6–10.0	6. 7	0.65	5.2–8.6
PHA-272-02/257-04	156	0.48	0.164	0.12–0.88	13.2	1.74	10.2–17.6	8.3	0.79	6.5–10.4	6.0	0.96	3.7–8.9
Inter-gene pool populations													
PHA-159-09/257-08	442	0.36	0.122	0.19–0.89	13.0	1.58	11.3–20.1	7.8	0.76	5.3–9.9	6.2	0.83	4.0–8.0
PHA-159-12/267-15	268	0.39	0.129	0.14–0.78	12.3	2.25	9.0–17.6	7.6	0.57	5.7–10.6	5.7	0.79	3.6–7.7
PHA-159-13/267-18	1095	0.41	0.145	0.09–0.91	13.0	1.84	8.5–18.3	7.7	0.98	4.6–10.0	5.7	0.73	1.4–8.4
PHA-159-14/257-14	393	0.30	0.102	0.11–0.77	11.7	1.42	7. 6–16.9	7.0	0.82	4.6–8.8	5.1	0.77	2.2–7.6
PHA-159-10/269-12	230	0.33	0.107	0.10–0.63	11.2	1.46	8.1–16.4	7.3	0.69	5.7–9.2	5.7	0.59	4.3–7.9
PHA-159-08/269-13	488	0.44	0.122	0.13–0.75	12.7	1.63	8.2–17.9	8.6	0.77	4.6–11.1	6.2	0.88	3.1–7.9

[a] Number of indivuals analyzed.

Table 4
Parental and progenie medians for seed traits of the intraracial, interracial and inter-gene pool common bean populations studied.

	P_1	P_2	F_2	F_3	F_4	F_5	F_6	F_7
Intraracial populations								
PHA-267-20/PHA-257-23 (NG × NG)								
N[a]	5	5		50	30	100	388	714
Weight	0.51	0.63		0.66**	0.77**	0.61**	0.60**	0.57**
Length	15.1	15.5		15.9**	17.5**	15.1**	14.9**	15.5
Height	7.5	8.0		8.6**	9.2**	7.9**	7.7**	8.0**
Width	6.1	6.7		6.8**	7.1**	6.6**	5.4**	6.4
Interracial populations								
PHA-452-01/PHA-119-01 (NG × P)								
N[b]	5	5			100	50		50
Weight	0.75	0.25			0.78**	0.73**		0.69**
Length	19.7	10.2			18.1**	17.1**		16.0**
Height	8.7	7.4			9.0**	8.2**		7.7**
Width	6.4	5.1			6.8**	6.6**		6.4**
PHA-257-01/PHA-323-02 (NG × CH)								
N[b]	5	5		50	67	86	100	125
Weight	0.54	0.70		0.88**	0.73**	0.70**	0.80**	0.80**
Length	14.3	17.2		17.8**	16.2**	16.2**	16.3**	17.6**
Height	8.8	8.4		9.2**	9.9**	8.0**	8.6**	9.3**
Width	7.6	7.1		7.8**	7.6**	6.9**	7.1**	7.5
PHA-257-06/PHA-306-01 (NG × P)								
N[b]	5	5		157	107	150	300	589
Weight	0.67	0.55		0.85**	1.10**	0.83**	0.76**	0.70*
Length	16.6	14.6		17.5**	19.2**	17.5**	16.9**	17.0**
Height	8.2	8.2		9.1**	9.0**	9.1**	8.4**	8.5**
Width	6.8	7.3		7.7**	7.7**	7.1**	6.9**	6.8**
PHA-267-18/PHA-338-27 (NG × P)								
N[b]	5	5		29	15		48	515
Weight	0.36	0.60		0.73**	0.76**		0.63**	0.64**
Length	15.3	14.7		16.3**	17.5**		15.8**	16.0**
Height	7.4	8.5		8.9**	9.8**		8.3**	8.5**
Width	6.0	7.1		6.5**	7.1**		6.8**	7.0**
Inter-gene pool populations								
PHA-159-09/PHA257-08 (M × NG)								
N[b]	5	5		135	75	72		140
Weight	0.24	0.58		0.35**	0.39**	0.35**		0.27**
Length	10.2	15.6		11.9**	12.9**	11.8**		11.8**
Height	6.9	7.8		7.2	7.9**	6.9**		6.6**
Width	4.5	6.4		5.2**	5.9**	4.8		4.9**
PHA-159-13/PHA-267-18 (M × NG)								
N[b]	5	5	11	200	150	275	189	250
Weight	0.27	0.56	0.15**	0.48**	0.48**	0.41**	0.35**	0.36**
Length	10.1	15.3	10.3**	14.0**	14.8**	11.7**	12.4**	12.9**
Height	7.1	7.5	6.5**	8.4**	8.7**	7.0**	7.7**	7.3**
Width	5.3	6.0	3.6**	6.0**	6.3**	5.3**	5.2**	6.0**
PHA-159-14/PHA-257-14 (M × NG)								
N[b]	5	5		100	50	65	58	150
Weight	0.30	0.51		0.36**	0.38**	0.26**	0.29**	0.25**
Length	12.0	15.1		12.1**	12.9**	11.3**	11.0**	10.1**
Height	7.2	8.0		7.3	8.0**	6.9**	7.1**	6.6**
Width	5.3	6.0		5.3**	5.7**	4.7**	5.0**	4.9**
PHA-159-08/PHA-269-13 (M × NG)								
N[b]	5	5	30	233	143			60
Weight	0.21	0.63	0.29**	0.48**	0.43**			0.38**
Length	9.7	16.1	12.2**	13.3**	12.1**			11.3**
Height	6.7	7.7	8.1**	8.9**	8.6**			8.3**
Width	4.5	6.7	5.6**	6.5**	6.2**			6.8**

Races—M: Mesoamerica; NG: Nueva Granada, P: Peru, CH: Chile.
[a] Number of indivuals analyzed.
* Generation significantly different at $P \geq 0.05$.
** Generation significantly different at $P \geq 0.01$.

18/PHA-338-27) were the best combinations because they displayed a positive transgression with respect to the progenitors for seed size and shape traits. Of the six inter-gene pool populations, only data was available for two F_2. One of them showed a hybrid depression (PHA-159-13/PHA-267-18), and the other one was skewed toward the Mesoamerican parent (PHA-159-08/PHA-269-13). Although inter-gene pool populations had an Andean parent with high seed weight (0.51–0.63), the mean seed weight of F_3–F_7 generations was significantly lower than that of mid-parent value.

Table 5
Segregation and chi-square goodness-of-fit tests for seed color pigmentation and protein markers in the F_3, F_4 and F_7 ($N \geq 30$) generations derived from intraracial, interracial and inter-gene pool common bean populations studied.

Marker	F_3[b] (0.375:0.250:0.375)		F_4[b] (0.437:0.125:0.437)		F_7[b] (0.492:0.0156:0.492)	
Intraracial populations						
PHA-20-07/PHA-159-11						
Pha: B:B/S:S	21:40:49	**	49:21:40		202:0:98	**
b11: b11/b13:b13	21:40:49	**	49:21:40		202:0:98	**
Me: 102:102/100:100	37:13:35		45:21:32		133:1:83	**
PHA-267-20/PHA-257-23						
Me: 100:100/98:98	[a]	[a]	[a]		155:0:347	**
Rbc: 100:100/98:98	[a]	[a]	[a]		63:0:439	**
Mdh2: 100:100/102:102	[a]	[a]	[a]		160:2:340	**
Interracial populations						
PHA-452-01/PHA-119-01						
C/W[c]	[a]		76:24	*	0:50	**
Pha: H:H/T:T	[a]		31:17:52	*	0:0:50	**
b6: b6/b7:b7	[a]		99:1:0	**	50:0:0	**
Me: 100:100/98:98	[a]		3:2:30	**	5:0:25	**
Rbc: 100:100/98:98	[a]		9:5:21		16:0:14	
Diap1: 100:100/95:95	[a]		11:1:23	*	7:0:23	*
Mdh1: 103:103/100:100	[a]		9:2:24	*	0:0:32	**
PHA-257-01/PHA-323-02						
C/W	2:48	**	40:27	**	0:100	**
Pha: T:T/C:C	48:0:2	**	25:5:38		86:0:0	**
b11: b11/b13:b13	48:0:2	**	24:6:37		86:0:0	**
Rbc: 100:100/98:98	19:0:22	*	2:5:23	**	32:0:50	
Mdh1: 103:103/100:100	16:4:1	**	27:3:0	**	64:0:18	**
PHA-257-06/PHA-306-01						
C/W	7:150	**	5:102	**	200:389	**
Pha: H:H/T:T	50:21:86	**	17:80:10	**	0:150:0	**
b6: b6/b7: b7	50:17:90	**	20:7:80	**	13:16:560	**
b11: b11/b12:b12	140:7:10	**	86:4:17	**	589:0:0	**
Me: 100:100/98:98	18:5:1	**	50:4:53		133:0:284	**
Rbc: 100:100/98:98	1:6:16	**	51:2:52		154:0:263	**
Mdh1: 103:103/100:100	10:7:6		69:8:30	**	417:0:0	**
PHA-257-10/PHA-306-11						
C/W	28:22	**	[a]		263:196	**
Pha: H:H/T:T	20:12:18		[a]		259:0:200	*
b11:b11/b12:b12	26:4:20		[a]		210:0:249	
Skdh: 103:103/100:100	19:5:19		[a]		11:0:374	**
Me: 100:100/98:98	13:4:16		[a]		87:0:298	**
Rbc: 100:100/98:98	15:2:16		[a]		62:0:323	**
Mdh1: 103:103/100:100	12:1:20	**	[a]		323:0:62	**
PHA-272-01/PHA-257-01						
C/W	44:6		21:9	**	45:25	**
Pha: H:H/T:T	13:32:5	**	12:5:13		70:0:0	**
b6:b6/b7:b7	26:8:16		16:2:12		0:0:70	**
b11:b11/b12:b12	17:13:20		14:1:13		0:0:70	**
Rbc: 100:100/98:98	11:2:17	*	15:0:15		38:0:32	
Inter-gene pool populations						
PHA-159-09/PHA-257-08						
Pha: S: S/T: T	10:125:0	**	0:74:1	**	0:140:0	**
b4: b4/b5: b5	39:9:87	**	22:3:50	**	140:0:0	**
b6:b6/b7:b7	12:2:121	**	30:0:45	**	90:0:50	**
b11:b11/b13:b13	60:29:46		46:6:23	**	140:0:0	**
Skdh: 103:103/100:100	24:16:56	**	14:6:51	**	39:0:62	*
Me: 100:100/98:98	22:15:59	**	7:5:59	**	0:0:101	**
Rbc: 100:100/98:98	33:3:60	**	47:2:22	**	61:0:40	
Diap1: 100:100/95:95	46:4:48	**	40:3:31	**	57:0:44	
Mdh1: 103:103/100:100	82:6:8	**	53:5:13	**	62:0:39	*
Mdh2: 102:102/100: 100	33:3:60	**	14:6:51	**	0:0:101	**
Inter-gene pool populations						
PHA-159-12/PHA-267-15						
Pha: B: B/T: T	[a]		39:1:10	**	49:50:49	**
b4: b4/b5: b5	[a]		8:2:40	**	99:0:49	**
b6:b6/b7:b7	[a]		10:1:39	**	49:0:99	**
b11:b11/b12: b12	[a]		48:2:0	**	49:0:99	**
Skdh: 103:103/100:100	[a]		24:2:4	**	45:0:72	*
Rbc[c]:100:100/98:98	[a]		21:3:6	**	48:0:69	
Diap1: 100:100/95:95	[a]		8:0:22	**	23:0:94	**
Mdh1: 103:103/100:100	[a]		8:1:21	**	49:50:49	**

Table 5 (*Continued*)

Marker	F_3[b] (0.375:0.250:0.375)		F_4[b] (0.437:0.125:0.437)		F_7[b] (0.492:0.0156:0.492)	
PHA-159-13/PHA-267-18						
Pha: S: S/T: T	27:165:8	**	122:13:15	**	0:250:0	**
b4: b4/b5: b5	36:85:79	**	45:33:72	**	196:4:50	**
b6:b6/b7:b7	87:28:85	**	81:11:58	*	50:0:200	**
b11:b11/b13:b13	60:94:46	**	53:65:32	**	93:9:148	**
Skdh: 103:103/100:100	39:15:26		62:1 0:15	**	127:0:82	**
Me: 102:102/100:100	31:11:38	*	49:9:29		138:0:71	**
Rbc: 100:100/98:98	21:6:53	**	20:15:52	**	20:0:189	**
Diap1: 100:100/95:95	3:9:68	**	5:6:76	**	82:0:127	**
Mdh1: 103:103/100:100	40:18:22		25:16:46	*	50:0:159	**
PHA-159-14/PHA-257-14						
Pha: S:S/T:T	1:99:0	**	1:49:0	**	50:125:0	**
b4:b4/b5:b5	1:0:99	**	1:49:0	**	48:127:0	**
b6:b6/b7:b7	19:50:31	**	18:15:17	**	65:4:106	**
b12:b12/b13:b13	46:7:47	**	46:0:4	**	75:0:75	
Skdh: 103:103/100:100	29:5:2	**	12:7:11		72:0:36	**
Me: 100:100/98:98	11:4:20	*	18:0:12		20:0:88	**
Rbc: 100:100/98:98	24:3:9	**	13:4:13		42:0:66	*
Mdh1: 103:103/100:100	27:0:9	**	24:0:6	**	66:0:42	*
Diap1:100:100/95:95	24:5:7	**	24:0:6	**	97:0:11	**
PHA-159-10/PHA-269-12						
C/W	[a]		49:11		0:30	**
Pha: S:S/T:T	[a]		21:24:15		30:0:0	**
b4:b4/b5:b5	[a]		16:15:29		0:0:30	**
b6:b6/b7:b7	[a]		16:10:34	*	0:0:30	**
b12:b12/b13:b13	[a]		34:26:0	**	30:0:0	**
Skdh: 103:103/100:100	[a]		19:0:32	**	20:0:6	*
Me: 102:102/98:98	[a]		45:0:6	**	26:0:0	**
Rbc: 100:100/98:98	[a]		24:4:23		40:7:43	
Diap1: 100:100/95:95	[a]		22:4:25		23:0:3	*
PHA-159-08/PHA-269-13						
C/W	208:1	**	120:23	**	59:0	
Pha: S:S/T:T	48:161:0	**	54:89:0	**	0:59:0	**
b4:b4/b5:b5	161:6:42	**	44:9:90	**	59:0:0	**
b6:b6/b7:b7	11:6:192	**	7:0:136	**	0:0:59	**
b12:b12/b13:b13	0:1:208	**	22:2:119	**	54:0:5	**
Skdh: 103:103/100:100	24:0:34	**	40:11:37		48:0:11	
Me: 102:102/98:98	15:0:43	**	32:9:47		4:0:23	**
Rbc: 100:100/98:98	32:5:21	**	36:9:43		24:3:0	**
Diap1: 100:100/95:95	39:0:19	**	20:1:68	**	0:0:27	**
Mdh1: 103:103/100:100	2:0:56	**	29:2:57	**	0:0:27	**

[a] Not available seed or with $N < 30$.
[b] Number of individuals in genotypic class. Mendelian ratios for codominant markers are showed in parentheses.
[c] C: color, W: White seeded. Color is a dominant marker, therefore, the Mendelian ratios 0.917:0.083 (F_3), 0.973:0.027 (F_4) and 0.999:0.001 (F_7) were used.
* Significant deviations from expected Mendelian ratios $P \geq 0.05$.
** Significant at $P \geq 0.01$.

3.3. Genetic marker analysis

The seed proteins studied were inherited in clusters as follows: b4/b5, b6/b7, b11/b12/b13 (lectins). Of the 8 isozymes loci examined, *Prx*^{98} and *Diap-2*^{105} did not reveal polymorphisms. Results of the tests for Mendelian ratios of the seed color pigmentation (3:1) and protein markers (1:2:1) of three generations are reported in Table 5. Even though F_3 and F_4 generations were not subjected to selection pressure, they showed high segregation distortion.

The selection increased the frequency of white-seeded genotypes in interracial populations and Mesoamerican genotypes in inter-gene populations. Seed protein markers showed higher proportion of skewed segregations (88% of the total polymorphic loci in the F_3) than the seed color pigmentation (67%) and isozymes (68%) markers.

Inter-gene pool populations showed a high proportion of skewed loci compared to intraracial and interracial populations. In intraracial populations, 50% of skewed loci were found in F_3 generation, while in interracial and inter-gene pool populations 67% and 87% were found, respectively.

Table 6
Frequencies of Phaseolin S and Phaseolin T alleles in the inter-gene pool populations.

Population	F3		F4		F7	
	S	*T*	*S*	*T*	*S*	*T*
PHA-159-09/PHA-257-08[a]	0.537	0.463	0.493	0.507	0.500	0.500
PHA-159-13/PHA-267-18	0.548	0.453	0.856	0.144	0.500	0.500
PHA-159-14/PHA-257-14	0.505	0.495	0.510	0.490	0.643	0.357
PHA-159-10/PHA-269-12	0.500	0.500	0.550	0.450	0	1
PHA-159-08/PHA-269-13	0.604	0.396	0.688	0.313	0.500	0.500

[a] The first parent corresponds with the Mesoamerican parent with Phaseolin S and the second with the Andean parent with Phaseolin T.

The analysis of allele marker frequencies in inter-gene populations showed a higher than expected frequency of alleles from the Mesoamerican gene pool (e.g. *Skdh*^{103} and *Diap-1*^{95} in the populations PHA-159-13/PHA-267-18 and PHA-159-08/PHA-269-13), many of which were fixed in the F_7 lines. Moreover, the superabundance of heterozygotes of phaseolin in these crosses was high even though the frequencies of the Pha (S)/Pha (B) and Pha (T) alleles tended to remain stable throughout the generations (Table 6).

4. Discussion

Information regarding the correlations among seed size traits is important for defining bean ideotypes for selection. Positive correlations among traits are desirable to make maximum genetic progress in seed size selection. In all crosses, large seed size was reflected by an increase in seed weight, height and length, but not always in seed width. Park et al. (2000) also reported high significant interactions between weight–length and weight–height, but not between length–width. This lack of correlation could be explained by the location of QTLs, four of the seven QTL for seed length and two of three QTL for seed height also appeared to correspond to QTL for seed weight (Park et al., 2000). The lack of association between seed length and width, and height and width in the inter-gene pool population PHA-159-12/PHA-267-15 could be the result of breakage of linkage or other genetic effects due to the distant origin of the parents.

The GCA provided an empirical estimate on the inheritance of the traits studied and it predicted a basis forecasting the value of the races used for future crosses. The significance of GCA indicated that these populations were different for frequencies of additive favorable alleles. Significant positive GCA effects were found for seed size in all large-seeded races, but the small-seeded Mesoamerica race showed significant negative GCA effects. The races Chile and Peru were the top combiners and might be desirable parents for improving seed size. Nueva Granada × Chile and Peru × Nueva Granada were the best specific combinations for seed size traits. These results indicated that genes associated with seed size in large-seeded parents combined positively in these crosses, because it was easier to fix these genes due to the additive nature. Small-seeded parents had significant negative effects on seed size, therefore may not be used to improve seed size.

The narrow genetic base of the intraracial populations may be a limitation to genetic gain (Beebe et al., 2000), which suggests that further progress in breeding for quantitative seed traits would be limited by a lack of genetic variation. Only the population PHA-267-20/PHA-257-23 within the Andean gene pool showed transgressive values in some seed traits (the average progeny values exceeded that of their parents and that of the parents with greater values) but with overall poor genetic progress.

The interracial populations are the next step before the hybridization between gene pools, so that there is no incompatibility caused by lethal genes (e.g. *DL1* and *DL2*) (Koinange and Gepts, 1992). The transgressive values in most of the interracial populations pointed out a significant gain from selection expressed as an increase in the seed size by means of the accumulation of large-seeded alleles. Other authors (Welsh et al., 1995; Santalla et al., 2005) found that progeny from interracial populations also showed the best values for agronomic traits. There seems to be useful genetic variation among landraces of Andean gene pool for seed size. Two combinations were notable for their transgressive values: White Kidney × Brown Garbanzo (PHA-257-01/PHA-323-02) and White kidney × Dark Garbanzo (PHA-257-10/PHA-306-11, PHA-257-06/PHA-306-01 and PHA-267-18/PHA-338-27). Genetic variation within Andean germplasm contributed to notable genetic progress in these populations by providing lines with a large seed size and excellent market potential. All large-seeded parents had positive GCA effects for seed size, indicating that larger-seeded populations contributed with favorable additive effects for seed weight to their progeny. Therefore, Andean races are potentially superior races for seed size, and should be included in breeding programs. Moreover, most large-seeded bean populations of this study have growth habit I or II. The need to mechanize and to reduce production costs and pesticide use in the production regions demands the development of upright cultivars with early-maturing bean. The bean cultivars should also combine good acceptance by consumers according to their seed quality. These populations were selected for seed commercial quality and seed yield from F_4 to F_7, so must be used in breeding programs as a source of useful traits for common bean improvement.

Inter-gene pool crosses generated large genetic variation in segregating populations, but most of the recombinants were of inferior performance (smaller seed size) to the parents themselves. Generally, the performance of the best line did not exceed that of the best parent. Thus, the behaviour of these populations is typical of what has been observed for the progeny of Andean × Mesoamerican crosses (e.g. Welsh et al., 1995; Santalla et al., 2005). Seed size is a highly heritable trait (narrow-sense heritability (h^2) estimates range from 0.54 to 0.80) controlled by several genes. Thus, this reduced performance may be due to the low probability of recovering genotypes with all large-seeded alleles necessary, given the population size and mating system used in this work. The wide genetic distance between the parents and the accumulation of some alleles could result in commercially undesirable phenotypes. These effects could not be seen until combined through crossing, which result in reduced values of the quantitative seed traits. Thus, despite of the opportunities that inter-gene pool crosses provide for increased levels of genetic variation, these recombinants have been less successful in comparison to the interracial populations. These results are in accordance with several studies that used American germplasm, which have recorded phenotypic abnormalities and a poor performance in the progeny of inter-gene pool populations (Singh and Gutièrrez, 1984; Gepts and Bliss, 1985; Welsh et al., 1995; Johnson and Gepts, 2002).

Crossability alone does not determine the success of introgression. Andean × Mesoamerican crosses have often not produced progeny superior to those crosses within the gene pools (e.g. Singh, 1995), perhaps due to the disruption of coadapted sets of genes. An alternative explanation could be that large-seeded recombinants were poor competitors and, hence, eliminated from the progeny. Thus, breeders must select early for large seed size. Furthermore, those using selection should form separate bulks according to seed size in early generations in populations showing a large segregation for the trait. Differences in genetic distance between gene pools and races dictate specific breeding methods and strategies. The frequency of useful genotypes recovered reduces with increasing genetic distance between parents, thus requiring a tailored approach to optimize the probability of success. Biparental crosses followed by pedigree, single seed descent or mass selections are poor methods for extracting adapted cultivars from Andean × Middle American populations. More elaborate programs of recurrent or congruity inbred-backcrossings (Urrea and Singh, 1995; Singh, 1999) are required to increase the probability of recovery of desirable homozygous progeny from the population. Other two related approaches have been proposed and used to increase the frequency of favorable QTL alleles at multiple loci: F2 enrichment followed by inbreeding and marker-assisted recurrent selection (MARS) (Bernardo, 2008).

Distorted segregation for molecular markers has been reported for several crops such as maize (*Zea mays* L.) (Sibov et al., 2003; Mano et al., 2005), barley (*Hordeum vulgare* L.) (Kleinhofs et al.,

1993; Cisthe et al., 2005), rice (*Oriza sativa* L.) (Xu et al., 1997; Lin et al., 1998) and common bean (Koenig and Gepts, 1989b; Vallejos et al., 1992, 2006; Paredes and Gepts, 1995). Segregation distortion in these crops has been attributed to gametic or sporophytic selection, incompatibility genes, and directed selection. The presence of a great percentage of markers that showed segregation distortion in inter-gene pool populations (87%) indicated that this phenomenon must be influenced by more distance related common bean genotypes.

The magnitude of the distortion became larger in advanced generations. The selection from F_4 was for seed commercial quality, which could have increased white-seeded genotypes in the interracial crosses, and for seed yield that could have favored the fixation of Mesoamerican genotypes in the inter-gene pool crosses.

A study carried out by Zamir and Tadmor (1986) in *Lens*, *Capsicum*, and *Lycopersicum* genera revealed that 54% of the loci in interspecific populations deviated from the expected segregation ratios, whereas 13% of the loci deviated from the expected ratios in intraspecific populations. In this study gene exchange in inter-gene pool populations involved parents that have evolved in separate ecological regions in response to different selection procedures, and these could differ at the molecular level. According to Johnson and Gepts (2002) the Andean and Mesoamerican common bean had been domesticated in different regions, and developed specific genic complex and epistatic interactions. When we crossed them, those combinations are dissolved, which reduces the likelihood of recovering promising lines.

Selfing of plants with differing allelic diversity results in different rates at which homozygosis can be reached. Lister and Dean (1993) reported 0.42% heterozygosis remaining at the F_8 generation in *A. thaliana*, while in an intraspecific tomato population, 2.1% heterozygosis remained in the F_6:F_7 generation (Saliba-Colombani et al., 2000) compared to 15% remaining in an interspecific population at F_7 (Paran et al., 1995). In this study, the highest recombination level was observed in inter-gene pool crosses. Three populations showed the Pha S/T heterozygote fixed at F_7 generation (PHA-159-09/PHA-257-08, PHA-159-13/PHA-267-18 and PHA-159-08/PHA-269-13) and 34% and 71% of heterozygosis remained at the F_7 generations in the populations PHA-159-12/PHA-267-15 and PHA-159-14/PHA-257-14, respectively. The process leading to heterozygote excess is not known but the characteristic is not uncommon in plant species (Liu et al., 1994; Paredes and Gepts, 1995; Kaló et al., 2000).

The small-seeded beans from Mesoamerica and large-seeded beans from Andes probably differ in key development pathways, which are disrupted upon intermating two contrasting parents. Some authors (Gepts and Bliss, 1985; Koinange and Gepts, 1992) have suggested the action of some form of female/male specific mechanism (nuclear–cytoplasmic interactions) that affects gene exchange between parental germplasm in inter-gene pool populations of common bean. The segregation distortion results in the preferential elimination of genetic material from one parent, in this case Andean parent, in hybrids plants. This distortion can compound the difficulties caused by low fertility because it will reduce opportunities for introgressive recombination.

Possible causes for this segregation distortion include genetic factors affecting the transmission of the genes. The consistency in the observed deviations in the form of heterozygotes excess suggests a possible genetic origin for diversity of alleles at the Phaseolin locus. The high percentage of heterozygotes for *Phs* locus detected could suggest that the Andean homozygous TT are not expressed in a Mesoamerican genetic background, since the Mesoamerican genotype was used as female parent, which would explain the lack of these homozygous. This appeared to be due to cytoplasmic × nuclear interactions or maternal effects, indicating that the direction in which a cross is made may have a perceptible effect on the progeny that be obtained from it. The underrepresentation of homozygous Andean genotypes could also represent the outcome of postzygotic elimination of genotypes possessing deleterious intergenomic combinations, providing an explanation for the embryo abortion often observed in this intergene crosses and the poor performance of its progeny, since this locus is associated with seed size.

The underrepresentation of Andean alleles, which have positive effects on seed weight, provided low gain for selection. Regions of segregation distortion resulted in an overabundance of genes in these regions from the Mesoamerican parent that had a negative phenotypic effect in seed size loci. Apparently, Mesoamerican alleles controlling small seed size were favoured above their Andean counterparts so that the large-seeded recombinants could not develop or were poor competitors and thus were eliminated from the Andean × Mesoamerican populations.

5. Conclusion and future prospects

Simultaneous selection for seed size and other important seed characteristics in inter-gene pool populations did not allow improving seed size. Large-seeded populations of Chile, Nueva Granada and Peru races showed better combining ability than their small-seeded Mesoamerican counterparts of Mesoamerica race. While significant seed size improvement was observed within Andean gene pool, inter-gene pool populations provided low genetic progress. Moreover, alleles from the Mesoamerican gene pool showed a higher frequency than expected, as possible result of its advantage competitive. Seed yield of large-seeded Andean cultivars of Chile and Nueva Granada races is from 40% to 60% lower than small-seeded counterparts from the Mesoamerican race. The results showed widespread deviations from Mendelian segregation and a significant degree of Andean genome was eliminated. The recombination ratio in such populations showed a strong distortion due to a high percentage of heterozygotes for *Phs* and other loci, indicating nuclear–cytoplasmic interactions that affects the development of the Andean homozigotes in Mesoamerican background. Congruity backcrossing between both gene pools should facilitate development of germplasm with intermediate characteristics. Subsequently, these could be used as parents for cultivar development.

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. The authors thank also to Diputación Provincial de Pontevedra (Spain) for farm facilities.

References

Bassiri, A., Adams, M.W., 1978. Evaluation of common bean cultivar relationships by means of isozyme electrophoretic patterns. Euphytica 27, 707–720.

Beebe, S.E., Skroch, P.W., Tohme, J., Duque, M.C., Pedraza, F., Nienhuis, J., 2000. Structure genetic diversity among common bean landraces of Mesoamerican origin based on correspondence analysis of RAPD. Crop Sci. 40 (1), 264–273.

Bernardo, R., 2008. Molecular markers and selection for complex traits in plants: learning from the last 20 years. Crop Sci. 48, 1649–1664.

Brown, J.W.S., Ma, Y., Bliss, F.A., Hall, T.C., 1981. Genetic variation in the subunits of globulin-1 storage protein of French bean. Theor. Appl. Genet. 59, 83–88.

Cisthe, L., Echavarri, B., Batlle, F., Moriano, M., Castillo, A., Valles, M.P., Romagosa, I., 2005. Segregation distortion for agronomic traits in doubled haploid lines of barley. Plant Breeding 124 (6), 546–550.

De la Cuadra, C., Martin, L., De la Rosa, L., Rodríguez, A., Varela, F., 2000. Trabajos actuales sobre *Phaseolus* desarrollados en el Centro de Recursos Fitogenéticos del INIA. In: Prólogo de las Actas de la Asociación Española de Leguminosas para el II Seminario de la Judía de la Península Ibérica, vol. 1. Villaviciosa, Spain, pp. 45–62.

Debouck, D.G., Toro, O., Paredes, O.M., Johnson, W.C., Gepts, P., 1993. Genetic diversity and ecological distribution of *Phaseolus vulgaris* (Fabaceae) in north-western South America. Econ Bot 47, 408–423.

Delaney, D.E., Bliss, F.A., 1991. Selection for increased percentage of phaseolin in common bean. I. Comparison of selection for seed protein alleles and S1 family recurrent selection. Theor. Appl. Genet. 81, 301–305.

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

Gepts, P., 1988a. Phaseolin as an evolutionary marker. In: Gepts, P. (Ed.), Genetic Resources of *Phaseolus* Beans. Kluwer, Dordrecht, The Netherlands, pp. 215–241.

Gepts, P., 1988b. A Middle American and Andean gene pool. In: Gepts, P. (Ed.), Genetic Resources of *Phaseolus* Beans. Kluwer, Dordrecht, The Netherlands, pp. 375–390.

Gepts, P., 1998. Development of an integrated genetic linkage map in common bean (*Phaseolus vulgaris* L.) and its use. In: Singh, S. (Ed.), Bean Breeding for the 21st Century. Kluwer, Dordrecht, The Netherlands, pp. 53–91 and 389–400.

Gepts, P., Bliss, F.A., 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., Bliss, F.A., 1986. Phaseolin variability among wild and cultivated common beans (*Phaseolus vulgaris*) from Colombia. Econ. Bot. 40 (4), 469–478.

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

Gepts, P., Nodari, R.O., Tsai, R., Koinange, E.M.K., Llaca, V., Gilbertson, R., Guzmán, P., 1993. Linkage mapping in common bean. Annu. Rept. Bean Improv. Coop. 36, 24–37.

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

Haley, S.D., Miklas, P.N., Afanador, L., Nelly, J.D., 1994. Random amplified polymorphic DNA (RAPD) marker variability between and within gene pools of common bean. J. Am. Soc. Hort. Sci. 119, 122–125.

Johnson, W.C., Menéndez, C., Nodari, R.O., Koinange, E.M.K., Singh, S.P., 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. Genom. 2 (previously J. Quant. Trait Loci).

Johnson, W.C., 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.

Kaló, P., Endre, G., Zimányi, L., Csanádi, G., Kiss, G.B., 2000. Construction of an improved linkage map of diploid alfalfa (*Medicago sativa*). Theor. Appl. Genet. 100, 641–657.

Kleinhofs, A., Kilian, A., Saghai Maroof, M.A., Biyashev, R.M., Hayes, P., et al., 1993. Molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. Theor. Appl. Genet. 86, 705–712.

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

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

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

Lin, S.Y., Sasaki, T., Yano, M., 1998. Mapping quantitative trait loci controlling seed dormancy and heading date in rice, *Oryza sativa* L., using backcross inbred lines. Theor. Appl. Genet. 96, 997–1003.

Lister, C., Dean, C., 1993. Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis*. Plant J. 4, 745–750.

Liu, C.J., Witcombe, J.R., Pittaway, T.S., Nash, M., Hash, C.T., et al., 1994. A RFLP-based genetic map of pearl millet (*Pennisetum glaucum*). Theor. Appl. Genet. 89, 481–487.

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

Mano, Y., Muraki, M., Fujimori, M., Takamizo, T., Kindiger, B., 2005. AFLP–SSR maps of maize teosinte and maize: comparison of map length and segregation distortion. Plant Breed. 124, 432–439.

Masi, P., 2001. Analysis of the genetic structure of *Phaseolus vulgaris* L. landraces in Basilicata, Italy. PhD dissertation, Università degli Studi della Basilicata, Potenza, Italy.

Nienhuis, J., Singh, S.P., 1988. Genetics of seed yield and its components in common bean (*Phaseolus vulgaris* L.) of Middle American origin. II. Genetics, variance, heritability and expected response from selection. Plant Breed. 101, 155–163.

Nodari, R.O., Koinange, E.M.K., Kelly, J.D., Gepts, P., 1992. Towards an integrated linkage map of common bean. I. Development of genomic DNA probes and levels of restriction fragment length polymorphism. Theor. Appl. Genet. 84, 186–192.

Osborn, T.C., Alexander, D.C., Sun, S.M., Cardona, C., Bliss, F.A., 1988. Insecticidal activity and lectin homology of arcelin seed protein. Science 240, 207–210.

Papa, R., Nanni, L., Sicard, D., Rau, D., Attene, G., 2005. The evolution of genetic diversity in *Phaseolus vulgaris* L. In: Motley, T.J., Zerega, N., Cross, H. (Eds.), Darwin's Harvest: new approaches to the origins, evolution, and conservation of crop. Columbia University Press, New York, USA, pp. 121–142.

Paran, I., Horowitz, M., Zamir, D., Wolf, S., 1995. Random amplified polymorphic DNA markers are useful for purity determination of tomato hybrids. HortScience 30 (2), 186–193.

Paredes, O.M., Gepts, P., 1995. Segregation and recombination in inter-gene pool crosses of *Phaseolus vulgaris* L. J. Heredity 86 (2), 98–106.

Park, S., Coyne, D.P., Jung, G., Skroch, P.W., Arnaud-Santana, E., Steadman, J., Ariyarathne, H., Nienhuis, J., 2000. Mapping of QTL for seed size and shape traits in common bean. J. Am. Soc. Hort. Sci. 125, 466–475.

Roach, D.A., Wulff, R.D., 1987. Maternal effects in plants. Ann. Rev. Ecol. Syst. 18, 209–235.

Rodiño, A.P., Santalla, M., González, A.M., De Ron, A.M., Singh, S.P., 2006. Novel genetic variation in common bean from the Iberian Peninsula. Crop Sci. 46, 2540–2546.

Saliba-Colombani, V., Causse, M., Gervais, L., Philouze, J., 2000. Efficiency of AFLP, RAPD and RFLP markers for the construction of an intraspecific map of the tomato genome. Genome 43, 29–40.

Santalla, M., Rodiño, A.P., De Ron, A.M., 2002. Allozyme evidence supporting south-western Europe as a secondary center of genetic diversity for common bean. Theor. Appl. Genet. 104, 934–944.

Santalla, M., Lema, M., Rodino, A.P., González, A.M., Monteagudo, A.B., De Ron, A.M., 2005. Improvement of large-seeded common bean cultivars under sustainable cropping system in Spain. Euphytica 142, 85–95.

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

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

Singh, S.P., 1995. Selection for seed yield in Middle American versus Andean × Mid-Middle American interracial common-bean populations. Plant Breed. 114, 269–271.

Singh, S.P., Gutierrez, J.A., 1984. Geographical distribution of the *DL1* and *DL2* genes causing hybrid dwarfism in *Phaseolus vulgaris* L., their association with seed size, and their significance to breeding. Euphytica 33, 337–345.

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

Singh, S.P., Gutierrez, J.A., Molina, A., Urrea, C.D., Gepts, P., 1991b. Genetic diversity in cultivated common bean II. Marker-based analysis of morphological and agronomic traits. Crop Sci. 31, 23–29.

Singh, S.P., Nodari, R.O., Gepts, P., 1991c. Genetic diversity in cultivated common bean I. Allozymes Crop Sci. 31, 19–23.

Singh, S.P., 1999. Integrated Genetic Improvement. In: Singh, S.P. (Ed.), Common Bean Improvement in the Twenty-First Century. Kluwer, Dordrecht, Netherlands, pp. 133–165.

Sibov, S.T., De Souza, C.L., Garcia, A.A., Garcia, A.F., Silva, A.R., Mangolin, C.A., Benchimol, L.L., De Souza, A.P., 2003. Molecular mapping in tropical maize (*Zea mays* L.) using microsatellite markers. 1. Map construction and localization of loci showing distorted segregation. Hereditas 139, 96–106.

Sprecher, S.L., 1988. Allozyme differentiation between gene pools in common bean (*Phaseolus vulgaris* L.) with special reference to Malawian germplasm. PhD thesis, Michigan State University, East Lansing/MI, USA.

Talbot, D., Adeng, M., Slightom, J., Hall, T.C., 1984. Size and organisation of multigene family encoding phaseolin, the major seed storage protein of *Phaseolus vulgaris* L. Mol. Gen. Genet. 198, 42–49.

Urrea, C.A., Singh, S.P., 1995. Comparison of recurrent and congruity backcrossing for interracial hybridization in common bean. Euphytica 81, 21–26.

Vallejos, C.E., Chase, C.D., 1991. Linkage between isozyme markers and a locus affecting seed size in *Phaseolus vulgaris* L. Theor. Appl. Genet. 81, 413–419.

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

Vallejos, C.E., Astua-Monge, G., Jones, V., Plyler, T.R., Sakiyama, N.S., Mackenzie, S.A., 2006. Genetic and molecular characterization of the I locus of *Phaseolus vulgaris*. Genetics 172 (2), 1229–1242.

Voysest, O., Dessert, M., 1991. Bean populations: classes and commercial seed types. In: Van Schoonhoven, A., Voysest, O. (Eds.), Common Beans: Research for Crop Improvement. C.A.B. Int., Wallingford, U.K. and CIAT, Cali, Colombia, pp. 119–162.

Weeden, N.F., 1984. Linkage between the gene coding the small subunit of ribulose bisphosphate carboxylase and the gene coding malic enzyme in *Phaseolus vulgaris*. Annu. Rept. Bean Improv. Coop. 27, 123–124.

Weeden, N.F., 1986. Genetic confirmation that the variation in the zymograms of 3 enzyme systems is produced by allelic polymorphism. Annu. Rept. Bean Improv. Coop. 29, 117–118.

Weeden, N.F., Liang, C.Y., 1985. Detection of a linkage between white flower color and *Est-2* in common bean. Annu. Rept. Bean Improv. Coop. 28, 87–88.

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

Xu, Y., Zhu, L., Xiao, J., Huang, N., Mccouch, S.R., 1997. Chromosomal regions associated with segregation distortion of molecular markers in F2, backcross, doubled haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). Mol. Gen. Genet. 253, 535–545.

Zamir, D., Tadmor, Y., 1986. Unequal segregation of nuclear genes in plants. Bot. Gaz. 147, 355–358.