Theor Appl Genet (2006) 112: 1149–1163
DOI 10.1007/s00122-006-0217-2

ORIGINAL PAPER

M. W. Blair · G. Iriarte · S. Beebe

QTL analysis of yield traits in an advanced backcross population derived from a cultivated Andean × wild common bean (*Phaseolus vulgaris* L.) cross

Received: 18 July 2005 / Accepted: 7 January 2006 / Published online: 24 January 2006

Abstract Advanced backcross QTL analysis was used to identify quantitative trait loci (QTL) for agronomic performance in a population of $BC_2F_{3:5}$ introgression lines created from the cross of a Colombian large red-seeded commercial cultivar, ICA Cerinza, and a wild common bean accession, G24404. A total of 157 lines were evaluated for phenological traits, plant architecture, seed weight, yield and yield components in replicated trials in three environments in Colombia and genotyped with microsatellite, SCAR, and phaseolin markers that were used to create a genetic map that covered all 11 linkage groups of the common bean genome with markers spaced at an average distance of every 10.4 cM. Segregation distortion was most significant in regions orthologous for a seed coat color locus (*R-C*) on linkage group b08 and two domestication syndrome genes, one on linkage group b01 at the determinacy (*fin*) locus and the other on linkage group b02 at the seed-shattering (*st*) locus. Composite interval mapping analysis identified a total of 41 significant QTL for the eight traits measured of which five for seed weight, two for days to flowering, and one for yield were consistent across two or more environments. QTL were located on every linkage group with b06 showing the greatest number of independent loci. A total of 13 QTL for plant height, yield and yield components along with a single QTL for seed size showed positive alleles from the wild parent while the remaining QTL showed positive alleles from the cultivated parent. Some QTL co-localized with regions that had previously been described to be important for these traits. Compensation was observed between greater pod and seed production and smaller seed size and may have resulted from QTL for these traits being linked or pleiotropic. Although wild beans have been used before to transfer biotic stress resistance traits, this study is the first to attempt to simultaneously obtain a higher yield potential from wild beans and to analyze this trait with single-copy markers. The wild accession was notable for being from a unique center of diversity and for contributing positive alleles for yield and other traits to the introgression lines showing the potential that advanced backcrossing has in common bean improvement.

Communicated by F. J. Muehlbauer

M. W. Blair (✉) · S. Beebe
CIAT—Centro Internacional de Agricultura Tropical,
Apartado Aéreo 6713, Cali, Colombia
E-mail: m.blair@cgiar.org
Tel.: +57-2-4450000
Fax: +57-2-4450073

*Present address:*G. Iriarte
CENICAFE—Centro Nacional de Investigación en Café,
Plan Alto Km 4 via antigua,
Chinchiná, Manizales, Colombia

M. W. Blair
CIAT—International Center for Tropical Agriculture,
1380 NW 78th Avenue, Miami, FL 33126, USA

Introduction

Common bean (*Phaseolus vulgaris* L.) is the most widely consumed grain legume worldwide and is highly appreciated in Latin America where it is part of the traditional diet (Broughton et al. 2003). The common bean cultigen was domesticated from wild *P. vulgaris*, a viny plant with indeterminate growth from the mid-altitude Neotropics and subtropics that has a wide distribution range from northern Argentina to northern Mexico (Gepts and Debouck 1991; Debouck and Smartt 1995). Domestications of cultivated common beans from diverging taxa of wild common beans are known to have occurred in two distinct centers of origin in South America and Central America, giving rise to the Andean and Mesoamerican gene pools, respectively (Gepts 1998; Chacón et al. 2005). The existence of the Andean and

Mesoamerican gene pools in both the wild and the cultivated gene pools is supported by diversity analysis for phaseolin seed proteins (Gepts and Bliss 1986; Gepts 1988), isozymes (Koenig and Gepts 1989; Singh et al. 1991b), and mitochondrial and genomic RFLPs (Becerra-Velazquez and Gepts 1994; Khairallah et al. 1992). The genetic diversity of cultivated common bean is thought to be smaller than that of wild common bean due to a genetic bottleneck and founder effect that occurred during crop domestication (Gepts and Bliss 1986; Gepts 1988; Koenig and Gepts 1989; Koenig et al. 1990; Cattan-Toupance et al. 1998; Chacón et al. 2005).

Within wild common beans, genetic diversity as measured by DNA based markers is large and spread over a wide range of North, Central, and South America. Freyre et al. (1996) studied 78 wild accessions with RAPD markers and found that they clustered into Andean and Mesoamerican gene pools as well as an intermediate gene pool of wild accessions from the Northern Andes. Meanwhile, Tohme et al. (1996) used AFLP markers to compare 114 accessions within a core collection of wild beans and found four gene pools among the germplasm—one from Mesoamerica and the other three from the Andean region. Among the wild common bean gene pools from South America, were two centered in the Northern Andes, one each in Colombia and Ecuador, and one centered in the Central and Southern Andes. Beebe et al (2001) studied 182 accessions of cultivated Andean beans compared to 29 accessions of wild beans and found that the cultivated Andean gene pool was likely to derive from the southern range of the wild bean distribution, namely from populations in Bolivia and northern Argentina. Furthermore, the cultivated Andean gene pool had relatively low diversity while the Northern Andean wild accessions from Ecuador and northern Peru were very distinct. Chacón et al. (2005) analyzed chloroplast DNA polymorphisms in 165 landrace accessions compared to 157 wild and weedy accessions and found evidence for only a single domestication in the Andean region compared to multiple domestications and wild × cultivated introgression events in Mesoamerica.

The genetic diversity results discussed above suggest that wild common beans are very diverse and can therefore be a useful source for enhancing the variability of cultivated common bean especially of the domesticated Andean gene pool which has low diversity. Several factors make wild common beans useful sources of diversity for incorporating novel and potentially useful characteristics into the cultigen: (1) wild beans do not have genetic barriers that prevent them from being crossed with cultivated beans; (2) a large range of ecotypes exist in wild beans from which to select; (3) wild beans have been subjected to natural selection which has resulted in potentially useful novel alleles; and (4) most wild alleles were not involved in domestication and remain untapped in the wild (Debouck and Tohme 1989; Kelly et al. 1999; Singh 2001; Gepts 2002). Gene transfer from wild to cultivated beans has been successful in one notable case: the transfer of monogenic Arcelin-based weevil resistance against the bruchid *Zabrotes subfasciatus* from a wild common bean to CIAT breeding lines (Kornegay et al. 1993). Singh et al. (1995) made the only attempt to use wild common beans for improvement of seed yield in a breeding program. These authors employed simple crosses and mass selection but had inconclusive results and recommended the use of backcrossing to assess the value of wild beans for yield improvement because of concurrent reductions in seed size that were observed in their progeny. Several authors (Singh et al. 1995; Gepts 2002) have pointed out some of the limitations to using wild common beans for breeding, such as (1) the undesirable agronomic characteristics of wild beans including poor architecture, small-seededness, and long growing period; (2) the practical difficulty in crossing wild beans to cultivated beans given their very different flowering and maturity regimes; and (3) the lack of breeding methodologies in the past to efficiently undertake wild × cultivated common bean crosses; as well as (4) the labor involved in crossing schemes to incorporate genes from wild beans into cultivated types.

Given these results a method is needed to overcome some of these barriers to greater use of the wild germplasm in common bean. One alternative is the advanced backcross method as suggested by Tanksley and Nelson (1996) that overcomes some of the limitations of using wild relatives in breeding programs. Advanced backcrossing is based on the inbred backcrossing that was effectively applied to common beans by Sullivan and Bliss (1983) to introgress seed protein content into bean cultivars from unadapted landraces. The advantage of these methods applied to wild × cultivated crosses is that they transfer favorable alleles from unadapted germplasm into elite breeding lines while avoiding the epistatic effects of deleterious genes found in the wild (Tanksley and McCouch 1997). A further advantage of the advanced backcross method is that quantitative trait locus (QTL) analysis of the resulting progeny can be used to identify positive alleles from the wild donor parent and these can be tracked in further crosses via marker-assisted selection (Tanksley and Nelson 1996). While the advanced backcross method has been applied to a large number of inbreeding crops especially among the cereals, the method has not been extensively applied to the legumes.

The objective of this research, therefore, was to test the advanced backcross method in common bean and to determine if positive QTL for yield and yield components could be detected within a $BC_2F_{3:5}$ population derived from a wild common bean (G24404) donor parent and an Andean ('ICA Cerinza') recurrent parent. The wild accession was notable for being from a unique center of diversity in the Northern Andes (Tohme et al. 1996), while the cultivar selected for improvement was a type I growth habit, large red-seeded, released variety in Colombia. The genetic map we developed was possible thanks to the recent development of co-dominant,

mapped microsatellite markers (Blair et al. 2003; Métais et al. 2002) a necessary prerequisite for this type of analysis. The study here is among the first to attempt to obtain a higher yield potential from wild beans and to analyze this trait with single-copy markers and we show that yield increasing alleles still reside untapped in wild accessions of common beans that can be exploited to improve cultivated types.

Materials and methods

Parental genotypes

The advanced backcross $BC_2F_{3:5}$ population was created from ICA Cerinza (the cultivated recurrent parent) crossed with G24404 (the wild donor parent). The recurrent parent, 'ICA Cerinza', henceforth called 'Cerinza', is a large red-seeded cultivar with type I growth habit that was released as a variety in 1991 and which is widely grown commercially in Colombia.

Cerinza was developed by G. Ligarreto at the Tibaitata experiment station of ICA (Instituto Colombiano Agropecuario) from the cross [(Antioquia 10 × L3043) × (Antioquia 8 × Antioquia 26)], where Antioquia 8, 10, and 26 are the Colombian landraces, Uribe Redondo, Algarrobo, and Sánchez, respectively (Voysest 2000). The variety belongs to the 'Radical' commercial class (with an average 100 seed weight of 53 g) and was widely tested in the departments of Cundinamarca, Boyacá, and Nariño before its release. The variety was found to be photoperiod insensitive and well adapted in a range of temperate to cool climates from 1,500 to 2,500 masl elevation in the Northern Andes and has been extensively adopted in Central and Southern parts of Colombia. Physiological maturity ranges from 75 days after planting in lower elevations to 140 dap in higher elevations. Cerinza has potential yields of 1,500 kg/ha on experiment station but yields about 600–1,000 kg/ha in farmers' fields. The variety is susceptible to ascochyta blight (*Ascochyta* spp.) but is resistant to anthracnose (*Colletotrichum lindemuthianum* Sacc.) and rust (*Uromyces phaseoli*); as well as tolerant to powdery mildew (*Erysiphe polygoni*), and root rots caused by *Fusarium* spp. Seeds of Cerinza have the 'T' phaseolin pattern that is typical of the Andean gene pool (this laboratory, unpublished results).

The donor parent, G24404, on the other hand, is a wild germplasm accession collected by O. Toro in 1992 from a site at 1,800 masl elevation in the Choachi municipality of Cundinamarca, Colombia. Like most wild beans, G24404 is a type IV climbing bean and produces small light and dark gray or black mottled seeds. It flowers irregularly usually at 95–110 days after planting when sown in September in Colombia, but matures over a long period of up to 120–350 days. The genotype has dehiscent pods with gray and black mottled, square-shaped seed with cream flecking and a 100 seed weight of 16 g. In a recent study, G24404 was found to have a mixture of phaseolin types including the 'CH' and 'L' patterns typical of Colombian wild and weedy accessions and the 'S' pattern which is typical of Mesoamerican genotypes (Chacón et al. 2005). Given this mixture, the accession was purified before its use in advanced backcrossing and only the wild 'CH' pattern plants were used for the crosses.

Population development

The advanced backcross $BC_2F_{3:5}$ population was developed in the following stages: (1) a simple cross was made between G24404 and Cerinza using the wild accession as a maternal parent and the cultivated genotype as a pollen source; (2) first backcrosses were made with six plants from the simple cross that were confirmed to be hybrids by morphological traits (intermediate growth habit and seed size); (3) second backcrosses were made on a plant-to-plant basis between 130 BC_1F_1 plants and the recurrent parent Cerinza used as a female parent. All hybridizations were performed in a screen house in Popayán, Colombia (18°C average temperature). The BC_2F_1 seed was harvested and planted in the field in the same location where it was advanced to the BC_2F_3 generation by single seed descent at which point the single plant selection was advanced by bulk harvesting for two generations. The $BC_2F_{3:5}$ generation was selected for two reasons: (1) two backcrosses were favored to return the majority of lines to the parental phenotype and (2) the F_3 generation of inbreeding was favored for the detection of additive genes. Generation advance to the $F_{3:5}$ generation also allowed for seed increase for the replicated trials. This was very important given that for Cerinza as for other Andean common beans there is a low multiplication ratio of approximately 30 seed harvested per seed planted due to its large seed size. The resulting $BC_2F_{3:5}$ population consisted of 157 lines, of which 95 were selected for bush architecture (type I, II, or III growth habits) and return to parental seed phenotype, while 62 lines were added without using any selection criteria to increase the chance for wild genome introgression.

Agronomic trials

The population was grown in three seasons at two locations: one season in Darién, Colombia (3°54′N, 76°30′W, 1,485 masl, average yearly temperature 20°C, average relative humidity 80%, average yearly rainfall 1,288 mm in 189 days of rain per year); and two seasons (A and B) in Popayán, Colombia (location 2°41′N, 76°66′W, 1,730 masl, average yearly temperature 18°C, average relative humidity 78.7%, average yearly rainfall 2,470 mm in 215 days of rain per year). All seasons had approximately 12-h photoperiod at flowering. The site in Darién has moderately acid, loam soils (pH 5.6, Andisol) while Popayán has acid, loam soils (pH 4.5,

Inceptisol). Darién and Popayán are both good growing environments for the Cerinza variety and therefore were appropriate testing sites for advanced backcross progeny derived from Cerinza; however, ascochyta blight is endemic in Popayán and control measures are needed for most bean production there. Given this, the experiment in the first season of Popayán (A) was protected against fungal diseases with four bi-weekly sprays of 2 l/ha of Bravo 500 (Chlorothalonil) fungicide (Syngenta) from the V4 growth stage onward while the experiment in the second season in Popayán (B) was protected with only two bi-weekly sprays before and after flowering and can be considered a disease treatment since ascochyta blight was aggressive in that season.

All other agronomic management was that recommended for dry beans in Colombia (fungicide seed dressing, preplanting application of superphosphate fertilizer, foliar applications of Bo–Zn chelates, manual weeding and hilling). The experiment in Darién included all 157 lines while the experiments in Popayán included the first set of 95 lines. Two row plots were 3 m long with a total of 12 seeds planted per meter and a distance between rows of 0.5 m. Each experiment consisted of a lattice design with three repetitions. Cerinza was used twice as a check variety in each repetition of each trial. G24404 was used as a visual check but not as a treatment in the trials given its very different architecture and phenology as well as its lack of adaptation to the second trial site in Darién.

Trait evaluation and data analysis

Yields were evaluated in all three seasons and were taken from the whole plot, eliminating the last two plants in each row to avoid border effects, and using harvested plant number as a covariable to estimate yield in kg/ha. Seed size, estimated from 100 seed weight in g and phenological traits, including days to flowering and days to harvest maturity were also evaluated in all three seasons. Physiological traits and yield component evaluated only in Darién included: plant height, plant width, seeds per plant, seeds per pod, and pods per plant. Plant height (in cm) was measured at the R6 growth stage from where the soil touched the hypocotyl to the furthest vegetative or reproductive meristem in indeterminate or determinate plants, respectively. Plant width (also in cm) was measured across the row taking into account leaf coverage at the R7 growth stage. Both measurements were made twice per plot and averaged prior to analysis. All quantitative data were analyzed using a general linear model and an analysis of variance in the software package SAS 8.2 (SAS Institute) and means were estimated to use for subsequent QTL analysis. Population distributions were evaluated for skewing, kurtosis, and normality (Lynch and Walsh 1998). Morphological characteristics, such as flower color, growth habit, seed color, and seed coat shininess that differed between the two parents were evaluated as qualitative data and confirmed over the two locations.

DNA extraction and molecular marker analysis

A bulk tissue sample of ten plants each was collected from seedlings of each line and parent grown in a greenhouse at CIAT and was used for DNA extraction following the method of Afanador and Hadley (1993). DNA quality was checked on a 0.8% agarose TAE gel, quantified fluorometrically, and diluted to a 10 ng/ul for use in PCR reactions. A parental survey was then carried out to identify polymorphisms with a total of 175 published microsatellite markers (Gaitán et al. 2002; Blair et al. 2003; Métais et al. 2002); four new cDNA derived microsatellites (BMc5, BMc9, BMc12, and VMd3) that are available from the corresponding author; and 12 SCARs (http://www.usda.prosser.wsu.edu/miklas/).

Microsatellite markers were amplified in a PTC-200 (MJ research) thermocycler using standard reagents (12.5 ng of genomic DNA, 2.5 mM of $MgCl_2$, 0.07 mM each dNTP, 0.05 mM each primer, 1 U *Taq* polymerase, and 1× reaction buffer in a 20 ul reaction volume) and conditions (5 min at 94°C, followed by 30 cycles of 1 min at 94°C, 1 min at 47°C, and 2 min at 72°C, followed by 5 min extension). Meanwhile, the SCAR markers were amplified with a PTC-100 (MJ research) thermocycler using a standard PCR reaction (50 ng of genomic DNA, 0.1 mM each primer, 0.2 mM dNTPs, 2.5 mM $MgCl_2$, 1 U *Taq* polymerase, 1× reaction buffer in a 25 ul volume) and annealing temperatures according to recommendations for each SCAR. Microsatellite markers were evaluated on silver-stained 6% polyacrylamide gels while SCAR markers were analyzed with ethidium bromide stained 2% agarose gels using standard laboratory techniques. A final molecular marker for the phaseolin seed protein was analyzed by SDS/PAGE gel electrophoresis according to recommendations by Gepts and Bliss (1986). The phaseolin pattern was compared to known standards provided by the Genetic Resource Unit of CIAT.

A total of 80 microsatellites, a single SCAR, and the phaseolin marker were used to evaluate the introgression level in the full set of lines. For each co-dominant marker, the individuals were coded as either homozygous for the donor parent allele, homozygous for the recurrent parent allele or heterozygous. For the dominant markers introgression events represented by the wild parent allele or band were scored.

Genetic mapping and QTL analysis

A genetic map was constructed using Mapmaker EXP v. 3.0 software (Lander et al. 1987) to perform linkage analysis (using the 'group' command) and to estimate the genetic distance between markers (using the 'map' command). Un-mapped markers were placed by using the 'try' command. Linkage group designations for each marker were checked against known microsatellite and morphological marker positions on other common bean genetic maps constructed by Blair et al. (2003) and

Freyre et al. (1998). Linkage group orientation and numbering are as reported in these previous references. Map distances are reported in centiMorgans (cM) estimated with the Kosambi mapping function. Markers were assigned with a minimum LOD value of 3.0 where LOD is the log 10 (L1/L0) and where L0 represents the probability of linkage and L1 represents the alternate hypothesis of no linkage. To approximate the $BC_2F_{3:5}$ generation, the B12 backcross population model of Mapmaker EXP v. 3.0 was used combining heterozygous and homozygous wild genotypic classes to represent the BC_2F_1 generation. The 'ripple' command also at a LOD of 3.0 was used to confirm marker order in each linkage group. Chi-square analysis was used to determine segregation distortion from the ratio of Cerinza and G24404 alleles expected for the $BC_2F_{3:5}$ generation.

Quantitative trait loci were identified using the genetic map constructed for the population, the phenotypic data, and the software program QTL Cartographer v. 2.0 (Basten et al. 2001). Significant QTL were found with model 6 composite interval mapping (CIM) analysis; a window size of 10 cM; and ten background markers used in a forward–backward stepwise multiple linear regression model. The CIM was done at every 1 cM (walkspeed) and the B12 genetic model of QTL Cartographer v. 2.0 was assumed. We used a default likelihood ratio (LR) threshold of 11.5 (Basten et al. 2001) to declare QTLs significant. In addition, we performed 1,000 permutations at a 5% significance level (Churchill and Doerge 1994) to identify the QTLs with the minimum LR of 11.5 that satisfied this stricter empirical threshold. Probability of a QTL locus was represented with a LR score where $LR = -2 \ln (L1/L0)$ and where L0 represents the probability of an association between the marker and the trait and L1 represents the alternate hypothesis of no association. LOD and LR are related by the formula $LOD = 0.2172LR$. The positions of the significant QTL were given for the maximum LR value within the region under analysis. The phenotypic variance controlled by a given QTL was determined by its determination coefficient (R2), while the phenotypic variance controlled by all the markers in the regression model was represented by a second determination coefficient (TR2) as defined by the software program. QTL were named based on Tar'an et al. (2002) who studied some of the same agronomic traits as well as a two-number code derived from the linkage group and the number of the QTL identified on that linkage group, separated by a period.

Results

Phenotypic variability in the advanced backcross population

Population histograms (Fig. 1) show that the yield was a quantitative trait with normal distribution in the advanced backcross population. Least significant difference comparisons showed significant variation between the introgression lines for yield and seed weight in all seasons. Average yields and yields of the best line per season were highest in Darién (1,393 and 2,245 kg/ha, respectively) followed by Popayán A (813 and 1,138 kg/ha), and Popayán B (236 and 510 kg/ha). Yields were lower in Popayán A and B due to endemic ascochyta blight disease pressure while Darién was not subject to the disease. In each season the population produced lines that were superior in yield by up to 30–250% over the recurrent parent, these being most significant in Popayán B season.

Values for 100 seed weight were normally distributed in both Darién and Popayán; however, the population averages in each season tended to be skewed towards smaller seed size (48–49 g per 100 seed) compared to the recurrent parent which had large (62 g per 100 seed), square-shaped grain in each season. Many of the introgression lines that out-performed Cerinza in terms of yield had smaller seeds than Cerinza but a few high yielding lines were of a similar size and shape to the recurrent parent. Introgression lines were also observed that had longer or rounder seed shape and darker red or purple seed coat color than the recurrent parent.

Phenological traits were also significantly different between lines in each season although average days to flowering and days to maturity of the introgression lines were close to those of the recurrent parent which flowers at 36–37 days after planting and matures at 78–83 days after planting in Darién and Popayán, compared to the wild accession which does not flower until 75 days after planting in Popayán and which does not mature regularly even at 114 days after planting at this location. Very few of the introgression lines were excessively late as this trait was selected against during development of the advanced backcross population; however, population distributions were significantly non-normal ($P = 0.01$) for days to maturity in both Popayán A and Darién; and for days to flowering in Darién. The non-normal population distributions for these traits suggest that days to flowering were associated with growth habit or some other discrete trait such as determinacy. While most of the introgression lines were determinate like the recurrent parent; a few (8 of 157) had indeterminate type II or III growth habit. Climbing ability had been selected against so there were no introgression lines with type IV growth habit.

Plant architecture traits measured in Darién also showed significant differences; however, the population distributions for these traits were skewed towards plants that were more compact than the recurrent parent (39 cm tall and 38 cm wide on average). These results were surprising but could be a result of introgression for smaller leaves and shorter floral racemes compared to Cerinza that has large leaves and long floral racemes, which make it a taller and wider plant type (45 cm tall and 44 cm wide on average). Significantly non-normal distributions ($P = 0.01$) were observed for both plant height and plant width.

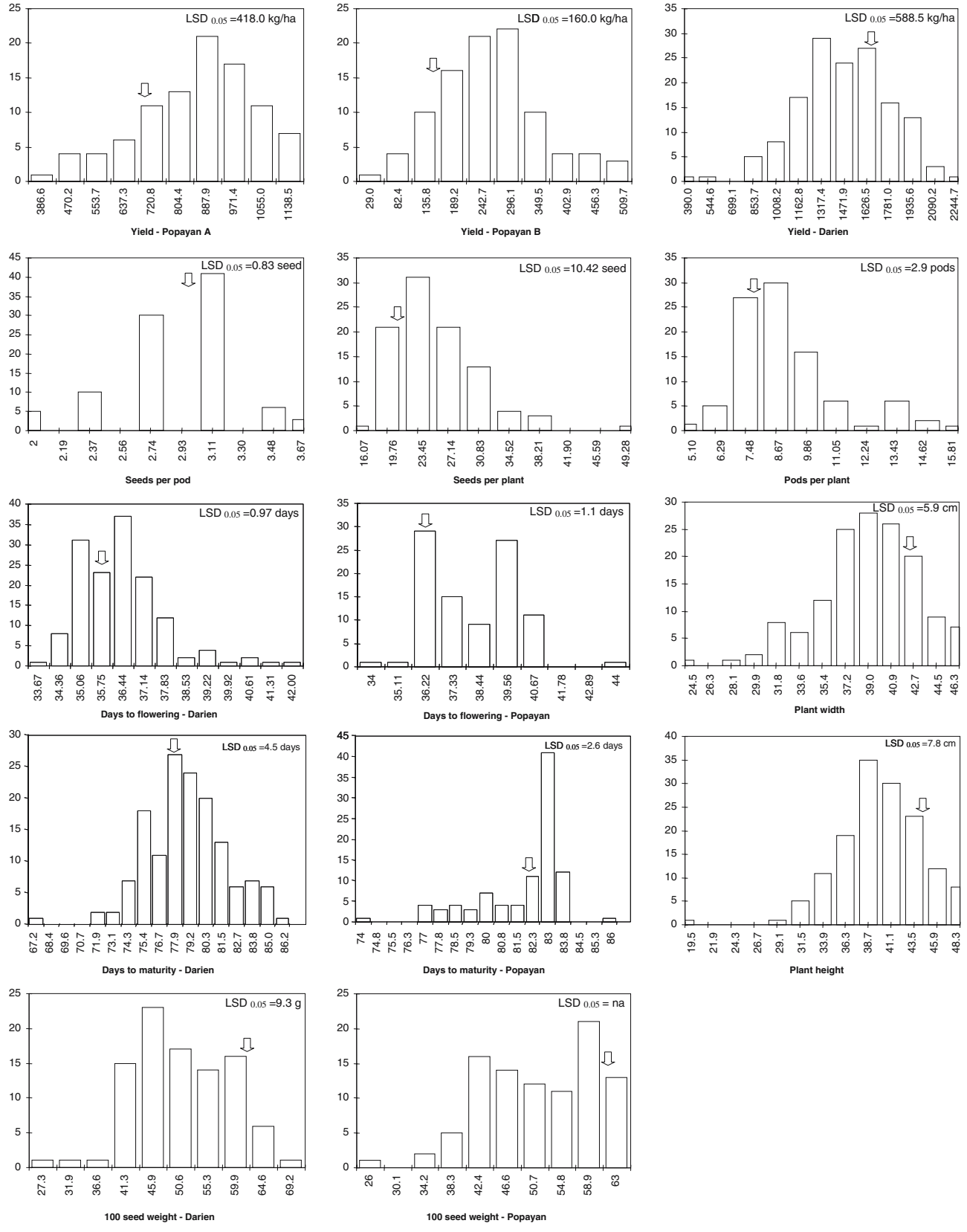

Fig. 1 Population distributions for yield, yield components (seeds per pod, seeds per plant, and pods per plant), phenological traits (days to flowering and days to maturity), plant architecture (plant width and plant height) and 100 seed weight in the {Cerinza × [Cerinza × (Cerinza × G24404)]} $BC_2F_{3:5}$ population**.** Least significant differences (LSD) are shown for each histogram and are based on the average of three replications. *Arrows* indicate phenotypic value of recurrent parent

For the yield component traits measured in Darién, the averages for the progenies (8.5 pods per plant, 2.8 seeds per pod, and 23.7 seeds per plant) were similar to those of the recurrent parent (8.1 pods per plant, 3.0 seeds per pod, and 23.2 seeds per plant); however, for both pods per plant and seeds per pod there were significantly superior progeny lines. Both pods per plant and seeds per pod were non-normally distributed ($P = 0.01$) while seeds per plant were distributed normally.

Polymorphism level and genetic mapping

The polymorphism level between the parents of the advanced backcross population was higher for microsatellite markers (45.7%) than for SCAR markers (41.6%). However, many of the SCAR markers were difficult to amplify with the wild accession DNA used as a control and only one was reliably mapped in the study. Among the microsatellites, genomic markers were more polymorphic than gene-based markers (data not shown). Polymorphisms were also identified between the parents in the phaseolin pattern and in morphological differences such as flower color and determinate/indeterminate growth habit. These traits were scored in the progeny lines to identify additional markers.

A genetic map was constructed for the advanced backcross population that covered all 11 linkage groups of the bean genome with a total genetic distance of 869.5 cM and a total of 84 markers including 80 microsatellites, one SCAR marker, two morphological markers and the phaseolin marker, all of which were single-copy markers that identified individual loci (Fig. 2). The average length of the linkage groups was 79.0 cM and there were 7.6 markers per linkage group. Marker distances in this map agreed with those of the microsatellite map generated by Blair et al. (2003) for the DOR364 × G19833 recombinant inbred line population showing that synteny between the maps is high. By extension the genetic map was syntenic with the genetic maps of Freyre et al. (1998), Nodari et al. (1993), and Vallejos et al. (1992) which are linked by single-copy microsatellite and RFLP markers to the map in Blair et al. (2003). Genetic distances in our study were smaller than these previous maps due to ours being a second backcross derived population. The largest linkage groups were b09 and b02, with 129.8 and 118.3 cM, respectively; while the smallest linkage groups were b08 and b01 with 18.2 and 33.9 cM, respectively. The linkage group with the greatest number of markers (12) was b02 while linkage group with the smallest number of markers (5) was b08. The only SCAR marker, SAP6, was placed between the microsatellite markers BM212 and BMc9 on linkage group b10, the first time that this marker has been definitively placed between other PCR-based markers.

Introgression rate and marker segregation

In the BC_2F_1 generation, 25% of plants were expected to contain an introgressed allele for a given locus, but all of these plants were expected to be heterozygous. Upon self-pollination, the loci were expected to segregate in a Mendelian ratio whereby in each generation heterozygosity was lost and the loci were fixed to homozygosity. In this study we analyzed BC_2F_3 derived families, in which the total amount of introgression from G24404 for a given locus was expected to be 15.675% of which, 9.375% was expected to be homozygous for the G24404 allele and 6.25% heterozygous for the G24404 allele. Nonintrogressed lines homozygous for the Cerinza allele were expected to represent 84.375% of the population for any given locus. Using the allele frequency for these expected segregation ratios and a significance threshold of $P < 0.01$, the chi-square test for the average amount of introgression per locus was not significantly different than expected across the majority of loci (59 out of 84 markers). However, the chi-square test at 25 loci differed significantly from the expected ratios described above and these were declared regions of segregation distortion and indicated by asterisks and plus signs next to the marker loci in the genetic map in Fig. 2 depending on the direction of segregation distortion and the probability level.

Segregation distortion was found to be most intense across almost all of linkage groups b01 and b08 as well as on parts of linkage group b02. Slight segregation distortion was observed on linkage groups b04 and b09, while a single marker on linkage groups b05 and b06 showed segregation distortion. Several linkage groups did not show any distorted markers. According to chi-square tests, the most distorted locus (BMd32) occurred on linkage group b01 while the least distorted loci (Pvat007 and BMd21) occurred on linkage group b09. Distortion was always against the wild G24404 allele and towards a predominance of the cultivated or Cerinza allele, except on linkage group b09 where there was segregation towards G24404 at BMd46 and BM154, and linkage group b07 where there was distortion towards G24404 near BM209, at the end of the linkage group. Despite the segregation distortion at one end of linkage group b07, segregation was normal near BM185 and the phaseolin locus. Since selection was made for return to recurrent parent plant architecture and seed type the predominance of the Cerinza allele at most loci was to be expected.

Significant quantitative trait loci

A total of 49 significant QTL were found across all traits and the three seasons but not for every trait × season combination as shown in Table 1. QTL found at the same genomic location in two or more seasons were considered a single QTL and named accordingly, reducing the number of named loci to a total of 41 QTL that are presented in Fig. 2. The QTL were spread over 11 linkage groups of the bean genome. The linkage

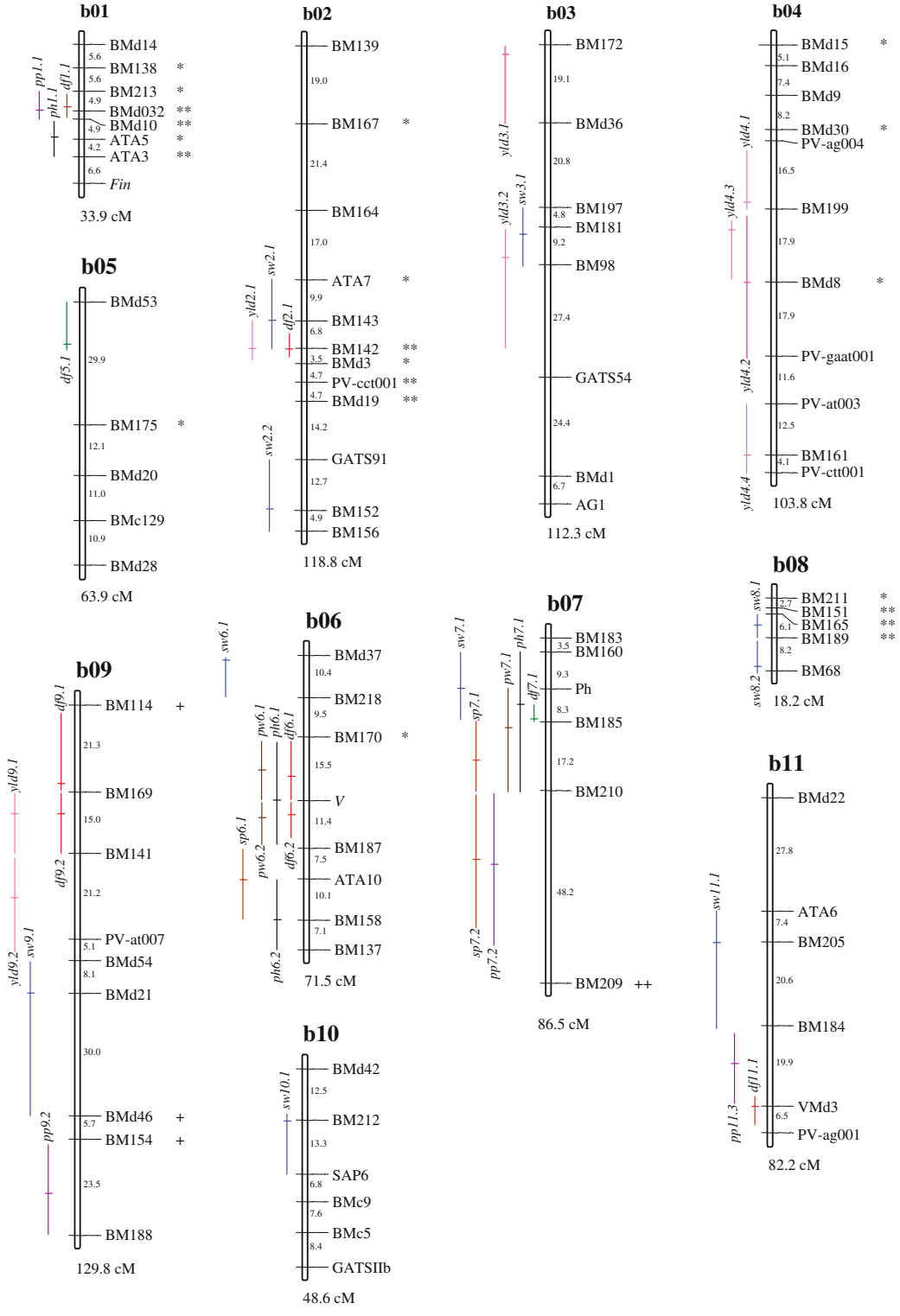

Fig. 2 Genetic linkage map for the advanced backross inbred line population developed from the cross {Cerinza × [Cerinza × (Cerinza × G24404)]} showing location of quantitative trait loci for yield, yield component, phenology, and plant architecture traits. Microsatellite (BM and PV) marker order was determined based on comparative mapping described in Blair et al. (2003) and LOD values of 3.0 or greater. Genetic distances are estimated in Kosambi cM units and linkage group designation is according to Freyre et al. (1998). QTL are identified as *vertical bars* crossed with a *horizontal bar* at the peak LR value. Abbreviations for traits are *df* days to flowering, *dm* days to maturity, *ph* plant height, *pp* pods per plant, *pw* plant width, *sp* seeds per plant, *sw* 100 seed weight, *yld* yield. Significant segregation distortions for genetic marker loci are indicated to the right of each linkage group for alleles from Cerinza (*, **) or from G24404 ($^{+}$, $^{++}$) with probability values of $P < 0.01$ and $P < 0.001$, respectively

Table 1 Quantitative trait loci for phenological traits, plant architecture, seed weight, yield and yield components identified by composite interval (CIM) mapping analysis for the Cerinza × G24404 advanced backcross population

QTL	Chr.	Nearest marker[a]	Flanking markers[b]	Location	Source	Increased effect	Significance values[c]		
							LR	R2	TR2
Days to flowering (days)									
*df*1.1	1	BMd32	BM213–BMd10	Darién	G24404	4.85	24.17	0.15	0.57
*df*2.1	2	BM142	BM143–BMd3	Darién	G24404	2.19	16.91	0.06	0.51
*df*6.1	6	BM170	BM170–*V*	Darién	G24404	2.37	19.53	0.16	0.61
*df*6.2	6	*V*	*V*–BM187	Darién	G24404	2.75	27.51	0.18	0.63
*df*9.1	9	BM114	BM114–BM169	Popayán A	G24404	2.36	**24.16**	**0.22**	**0.38**
	9	BM114	BM114–BM169	Darién	G24404	1.46	26.67	0.13	0.55
*df*9.2	9	BM169	BM169–BM141	Popayán A	G24404	2.41	**22.20**	**0.22**	**0.39**
	9	BM169	BM169–BM141	Darién	G24404	1.67	**32.93**	**0.16**	**0.58**
*df*11.1	11	VMd3	BM184–PVag001	Popayán A	G24404	2.15	**14.37**	**0.11**	**0.36**
Days to maturity (days)									
*dm*5.1	5	BMd53	BMd53–BM175	Popayán A	Cerinza	4.08	14.55	0.37	0.72
*dm*7.1	7	*Ph*	*Ph*	Popayán A	Cerinza	3.93	13.51	0.14	0.40
Plant height (cm)									
*ph*1.1	1	ATA5	BMd10–ATA3	Darién	G24404	9.38	**20.66**	**0.09**	**0.44**
*ph*6.1	6	*V*	BM170–BM187	Darién	Cerinza	6.70	**23.71**	**0.10**	**0.44**
*ph*6.2	6	BM158	ATA10–BM137	Darién	G24404	6.09	**19.19**	**0.08**	**0.44**
*ph*7.1	7	Ph	BM160–BM210	Darién	Cerinza	5.82	**26.13**	**0.13**	**0.44**
Plant width (cm)									
*pw*6.1	6	BM187	BM170–*V*	Darién	Cerinza	5.56	**19.98**	**0.17**	**0.30**
*pw*6.2	6	*V*	*V*–BM187	Darién	Cerinza	4.92	**17.89**	**0.13**	**0.31**
*pw*7.1	7	BM185	*Ph*–BM210	Darién	Cerinza	6.52	**21.01**	**0.19**	**0.36**
Yield (kg/ha)									
*yld*2.1	2	BM142	BM143–BMd3	Popayán B	G24404	225.04	**16.56**	**0.09**	**0.52**
*yld*3.1	3	BM172	BM132–BMd36	Popayán A	Cerinza	219.41	12.43	0.10	0.50
*yld*3.2	3	BM98	BM181–GATS54	Popayán B	G24404	98.57	**17.37**	**0.11**	**0.54**
*yld*4.1	4	PVag004	PVga004–BM199	Popayán B	Cerinza	152.63	**18.36**	**0.13**	**0.56**
*yld*4.2	4	BMd8	BM199–PVgaa001	Popayán A	Cerinza	277.78	**21.56**	**0.14**	**0.46**
*yld*4.3	4	BM199	BM199–BMd8	Popayán B	Cerinza	150.91	**18.10**	**0.13**	**0.56**
*yld*4.4	4	BM161	PVat003–PVctt001	Popayán B	G24404	219.94	**34.79**	**0.21**	**0.52**
*yld*9.1	9	BM169	BM169–BM141	Darién	Cerinza	326.24	**13.64**	**0.11**	**0.30**
*yld*9.2	9	BM141	BM141–BMd54	Popayán A	G24404	179.75	12.81	0.12	0.52
	9	BM141	BM141–PVat007	Popayán B	G24404	104.30	12.11	0.12	0.60
Pods per plant									
*pp*7.2	7	BM210	BM210–BM209	Darién	G24404	4.95	**39.38**	**0.64**	**0.65**
*pp*9.2	9	BM154	BM154–BM188	Darién	G24404	4.98	**40.04**	**0.63**	**0.64**
*pp*11.3	11	BM184	BM184–VMd3	Darién	G24404	5.22	**44.79**	**0.64**	**0.65**
Seeds per plant									
*sp*6.1	6	ATA10	BM187–BM158	Darién	G24404	12.38	20.38	0.15	0.37
*sp*7.1	7	BM185	BM185–BM210	Darién	G24404	7.46	13.60	0.16	0.45
*sp*7.2	7	BM210	BM210–BM209	Darién	G24404	8.06	15.09	0.29	0.57
100 seed weight (g)									
*sw*2.1	2	BM143	ATA7–BM142	Popayán A	Cerinza	6.11	14.50	0.04	0.77
*sw*2.2	2	BM152	GATS91–BM156	Popayán A	Cerinza	9.81	**44.13**	**0.16**	**0.79**
	2	BM152	GATS91–BM156	Darién	Cerinza	10.24	**45.52**	**0.17**	**0.72**
*sw*3.1	3	BM181	BM197–BM98	Popayán A	Cerinza	6.80	**21.92**	**0.07**	**0.78**
*sw*6.1	6	BMd37	BMd37–BM218	Darién	G24404	9.62	**15.25**	**0.06**	**0.74**
*sw*7.1	7	*Ph*	BM160–BM185	Popayán A	Cerinza	8.35	**22.70**	**0.07**	**0.75**
	7	*Ph*	BM160–BM185	Darién	Cerinza	7.64	**19.37**	**0.06**	**0.73**
*sw*8.1	8	BM165	BM165–BM189	Popayán A	Cerinza	11.40	14.20	0.05	0.75
	8	BM165	BM165–BM189	Darién	Cerinza	12.14	17.23	0.08	0.67
*sw*8.2	8	BM189	BM189–BM68	Popayán A	Cerinza	9.06	26.29	0.08	0.77
	8	BM189	BM189–BM68	Darién	Cerinza	7.97	18.30	0.06	0.73
*sw*9.1	9	BMd21	BMd54–BMd46	Darién	Cerinza	6.32	18.37	0.07	0.74
*sw*10.1	10	BM212	BMd42–SAP6	Darién	Cerinza	7.14	13.81	0.04	0.73
*sw*11.1	11	BM205	ATA6–BM184	Popayán A	Cerinza	10.28	**23.41**	**0.08**	**0.79**
	11	BM205	ATA6–BM184	Darién	Cerinza	7.64	**25.41**	**0.09**	**0.73**

Values in **bold** indicate significance at 0.5% probability after 1000-fold permutation tests. Values in nonbold indicate significance at LR 11.5 or above

[a]Nearest marker is the marker closest to the peak LR score

[b]Markers on each side of the significant QTL are indicated as flanking markers

[c]*LR* likelihood ratio test statistic for H_0:H_1 where H_0 is the hypothesis of no QTL effect at test position and H_1 is the hypothesis of a QTL effect at the test position; *R2* proportion of variance explained by the QTL at test site; *TR2* total R2 for the QTL and the background markers

group with the greatest number of QTL was b06 with eight significant QTL. Other linkage groups with large numbers of QTL were b07 and b09 with seven and six QTL, respectively. Linkage groups b02 and b04 each had four QTL, linkage groups b03 and b11 each had three QTL, and linkage groups b01 and b08 each had two QTL. Linkage groups b05 and b10 had the fewest QTL with only one each.

Phenological traits

A total of nine QTL were found for days to flowering and days to maturity across the two locations. QTL for days to flowering in Darién were found on linkage groups b01 at BMd32, b02 at BM142, and b06 in the interval between BM170 and BM187 that flanked the *V* gene for flower color. Two QTL on linkage group b09 in the interval between BM114 and BM141 with peaks near BM169 were significant in both Darién and Popayán. A final QTL for days to flowering on linkage group b11 was only significant in Popayán, as were two QTL for days to maturity on linkage groups b05 near Bmd53 and b07 near the Phaseolin locus. The most significant QTL for days to flowering were *df9.1* and *df9.2* on linkage group b09 that explained 13–22% of the variance each and were associated with the late flowering allele from G24404. Meanwhile, the most significant QTL for days to maturity was *dm5.1* on linkage group b05 that explained 37% of variance for the trait in Popayán and was associated with the Cerinza allele. Surprisingly, there was no overlap between the QTL for days to flowering and those for days to maturity.

Plant architecture

A total of seven QTL were found for plant height and width in Darién. QTL for plant height were found on linkage group b01 at the ATA5 locus, on b06 at separate locations near the *V* locus and near BM158, and on b07 at the phaseolin locus. The first two QTL were associated with the QTL described above for days to flowering that were located on b01 and b06. QTL for plant width were found at the same location as QTL for plant height on the b06 linkage group near the *V* locus and on the b07 linkage group near the phaseolin locus. The QTL for plant height explained from 8 to 13% of variance while the QTL for plant width explained 13–19% of variance. Greater plant height was associated with the G24404 allele at *ph1.1* and *ph6.2*; while greater plant height and width were associated with the Cerinza allele at the remaining QTL on linkage group b06 (*ph6.1*/*pw6.1*/*pw6.2*) and linkage group b07 (*ph7.1*/*pw7.1*).

Yield

A total of nine QTL were found for yield over the two locations. QTL for yield in Popayán were found on linkage group b02 near BM142 for the second seasons' data (*yld2.1*), on b03 at two separate locations for both seasons at this location (*yld3.1* and *yld3.2*), on b04 across three regions of the linkage group depending on the season (*yld4.1*, *yld4.2*, *yld4.3*, and *yld4.4*), and on b09 near BM141 for both seasons (*yld9.2*). A single QTL for yield in Darién (*yld9.1*) was found on linkage group b09 at an adjacent interval to *yld9.2* near BM169. The positive allele for the yield QTL varied depending on the locus identified, with five QTL showing a positive effect of the G24404 allele and four QTL showing a positive effect of the Cerinza allele. The positive effect of the QTL on linkage group b02 was from G24404 while those for each of the linkage groups with two QTL (b03 and b09) were from both parents. The pattern of QTL on linkage group b04 was interesting because all the proximal QTL (*yld4.1*, *yld4.2*, and *yld4.3*) had positive effects from Cerinza over both seasons while the terminal QTL (*yld4.4*) had a positive effect from G24404 in the second season. The QTL for yield explained from 9 to 21% of the variance for this trait in these seasons and had phenotypic effects that ranged from 98 to 326 kg/ha in increased or decreased yield. It was notable that the *yld9*.1 QTL from G24404 was consistent across two seasons in Popayán and was associated with the adjacent *yld9.2* QTL from Cerinza. Significant genotype × environment interaction was also evident for the QTL *yld2.1* and *yld3.2* that were both found in Popayán B, the season with the lowest yields, but were not detected with yield data from any other season. This could be explained by the serious infestation of ascochyta blight during this season, indicating that these yield QTL may be for vigor under disease pressure or disease tolerance per se rather than for yield itself. Meanwhile QTL *yld3.1*, *yld4.3*, *yld9.1*, and *yld9.2* were found in seasons without disease pressure and may represent yield QTL associated with other factors.

Yield components

A total of six QTL were identified for yield components in Darién, three each for pods per plant and number of seeds per plant that surpassed threshold LR values of 39.2 and 24.1, respectively. No QTL for seed per pod were identified using a threshold LR value of 13.8. The QTL for number of pods per plant were found on linkage groups b07, b09, and b11. The variance explained by the QTL reached 64%, while the positive effect on number of pods per plant was always from the wild parent G24404 allele with a phenotypic effect that varied from 4.9 to 5.2 more pods per plant. One QTL for seeds per plant was located on linkage group b06 near ATA10 while two others were located on b07 at two separate locations near BM210. Similarly to results with pods per plant, the QTL for number of seeds per plant had positive effects from the G24404 allele although the variance explained ranged from 15 to 29%. The phenotypic effect of these QTL ranged from 7.5 to 12.4

more seeds per plant. The QTL *sp7.2*. for seeds per plant overlapped in location with the QTL for pods per plant *pp7.2*.

Seed weight

A total of 11 QTL were identified across eight linkage groups for 100 seed weight at the two locations. The QTL *sw2.2*, *sw7.1*, *sw8.1*, *sw8.2*, and *sw11.1* were consistent across both Darién and Popayán while the remaining QTL were specific to one location or the other and evenly split between these two locations. The positive allele for seed size came from the larger seeded cultivated parent Cerinza in all cases except for the *sw6.1* QTL that was identified in Darién and came from the small seeded wild parent G24404. The phenotypic effect of the QTL on 100 seed weight varied from 6.0 to 12.1 g, while the variance explained by individual QTL was 4–17% due to the large number of QTL identified for the trait. All the QTL were found linked to microsatellite markers except for the QTL on linkage group *sw7.1* which was found linked to the phaseolin locus. Seed weight QTL on linkage group b03 and b09 were linked to QTL for total seed yield while QTL for seed weight on b09 and b11 were linked with QTL for number of pods per plant. The QTL for seed weight on linkage groups b08 and b10 were not associated with QTL for other traits.

Discussion

Usefulness of the genetic map

Microsatellite markers were ideal for this study because they had the advantages of being (1) co-dominant and therefore able to diagnose the introgression lines derived from heterozygous versus homozygous BC_2F_1 genotypes; (2) highly polymorphic so that most microsatellites evaluated could be used for genetic mapping; and (3) single copy and therefore useful for genetic map construction. The final genetic map contained a marker ever 10.4 cM and had close to full genome coverage for all 11 chromosomes of common bean that made it ideal for QTL detection using CIM. Given that most of the microsatellites had been previously positioned on two recombinant inbred line genetic maps (Blair et al. 2003; Yu et al. 2000), it was possible to make comparisons that showed synteny between the advanced backcross map and previous genetic maps based mostly on recombinant inbred line populations. Due to the high polymorphism found between the wild and cultivated parents, the genetic map for the advanced backcross population also proved to be a useful resource for mapping microsatellites from Gaitán et al. (2002), Métais et al. (2002), and Blair et al. (2003) that were monomorphic or untested in the genetic mapping studies discussed above.

The advanced backcross population was also useful for placing two morphological markers: (1) for flower color which was mapped based on segregation of purple and white flower color alleles inherited from G24404 and Cerinza, respectively, and which was found to be located on linkage group b06 in a position equivalent to the *V* locus mapped by Nodari et al. (1993); and (2) for determinate growth habit which was mapped based on the segregation of alleles for determinate and indeterminate terminal axes inherited from Cerinza and G24404, respectively, and which was found to be located on linkage group b01 in a position equivalent to the *fin* gene for determinacy (Coyne 1967, 1970) that was mapped by Koinange et al. (1996). Finally, the advanced backcross population was in agreement with previous reports for the placement of the phaseolin locus on linkage group b07 (Vallejos et al. 1992; Nodari et al. 1993) and the SAP6 SCAR marker on linkage group b10 (Miklas et al. 2003).

Segregation distortion

Segregation distortion was a significant feature of the advanced backcross population and chi-square tests showed that for many loci the introgression was below that expected based on Mendelian segregation (10 markers at $P < 0.001$ and 15 markers at $P < 0.01$ as shown in Fig. 2). Segregation distortion is a common feature of inter-gene pool crosses in common bean (Freyre et al. 1998; Johnson and Gepts 2002) and has also been found in a balanced population created from a wild × cultivated bean cross (Koinange et al. 1996). Segregation distortion in other crops has been attributed to gametic or sporophytic selection, incompatibility genes, and directed selection (Xu et al. 1997). Regions of segregation distortion probably reflected the presence of genes in these regions from the wild parent that had a negative phenotypic effect and that were selected against along with linkage drag around these loci.

Indeed, several of the regions of the genome that were under-represented and possibly selected against in the development of the advanced backcross population coincided with domestication syndrome genes that have been mapped previously in common beans (Koinange et al. 1996) and which were known to be selected against during the development of beans as a modern crop including such traits from the wild as indeterminacy, pod dehiscence, and small seededness (Singh 2001; Gepts 2002). High segregation distortion and lack of introgression on linkage group b01 would have been the result of the elimination of advanced backcross lines with this wild introgression and linkage drag with the *Fin* and *Ppd* genes that are known to be located on this linkage group (Koinange et al. 1996; Gu et al. 1998). Wild alleles at these loci would condition indeterminate growth habit and photoperiod sensitivity and both of these traits would have been agronomically negative factors that were selected against in this cross, since the

desired plant type was an early flowering, bush bean like the cultivated recurrent parent.

Selection also caused segregation distortion on linkage group b02 which has been shown to contain QTL for seed dormancy near the *St* gene for pod fiber and seed shattering, a trait that is common in wild beans but selected against in cultivated common beans (Koinange et al. 1996). Meanwhile, segregation distortion at four marker loci on linkage group b08 could have been due to selection for return to the recurrent parent seed type with fixation of the [*R-C*] locus which determines seed coat pattern and color intensity and has been mapped to that location of the genome (McClean et al. 2002). A cluster of domestication syndrome trait QTL for plant architecture and photoperiod sensitivity were also identified on this linkage group by Koinange et al. (1996) and also could have been selected against in this study in the return to recurrent parent growth habit and phenology as mentioned above. Less severe segregation distortion on linkage groups b04 and b06 could have been due to selection of other pod traits or seed color genes and modifiers, including potentially the *Z* gene for partial seed coat which is linked with the *fin* gene on linkage group b01 according to Bassett (1991). The process of advanced backcrossing and selection can be considered to have replicated what occurred at the level of domestication with the elimination of domestication syndrome alleles from the wild with selective sweeps occurring around the genes that were critical to traits of interest to early farmers during domestication. Many of these traits are likely to have been associated with both agro-ecological adaptation as well as appearance or seed coat color.

Segregation distortion and selection against the wild parent alleles did not affect the discovery of QTL as these were detected even in regions where two or more adjacent markers were subject to negative selection such as those on linkage groups b01, b02, and b08. Meanwhile, selection for positive genes from the wild parent was evident on linkage groups b04 and b09. The reason for such selection would be largely unknown as positive genes from wild accession have been less well studied than negative genes from wild accession (Singh et al. 1995; Koinange et al. 1996) but could be associated with positive qualitative or quantitative traits that were identified in this study in terms of yield or yield components. It was surprising that in this study segregation distortion was not greater around the phaseolin locus on linkage group b07 which is known to be associated with seed size in the domestication syndrome (Koinange et al. 1996) and in Mesoamerican × Andean crosses (Anthony et al. 1990; Kami et al. 1995). The large difference in seed size between the cultivated and the wild parents seems to be controlled by a larger set of genes than the difference in seed size between Andean and Mesoamerican cultivars as seen by the large number of QTL affecting this trait in the advanced backcross population. These results agree with the QTL studies by Koinange et al. (1996) who found seed size loci on linkage groups b01 and b07 as well as Park et al. (2000) who found that seed size was controlled by loci on linkage groups b03 through b06 in addition to b07. Similar to our study, Tar'an et al. (2002) identified QTL for seed size on linkage groups other than b07 (namely b04 and b11) in their evaluation of a cross of two small-seeded Mesoamerican genotypes.

Co-localization and consistency of QTLs

A range of QTL were detected across more than one environment in this study with good consistency for days to flowering QTL on linkage group b09, for seed weight QTL on linkage groups b02, b07, b08, and b11, and yield QTL on linkage group b04 and b09 (Table 1). The consistency of many of the seed weight QTL may reflect the higher heritability of this trait even though QTL analysis suggests the trait is controlled by multiple loci in crosses between wild and cultivated common beans. Several of the QTL for higher yield or more pods per plant were associated with QTL for smaller seed weight or later flowering. These negative associations could be due to pleiotropy or linkage of QTL for these traits. The yield QTL on linkage group b04 were a special case as they were from the wild parent and were not associated with yield components, seed size or flowering time. Since these QTL were consistent, found in a chromosomal region of known resistance genes (Kelly et al. 2003) and resistance gene analogs (López et al. 2003) and occurred for the season where ascochyta blight was significant, some of the yield differences could have been due to tolerance to this disease rather than yield potential itself. A resistance QTL for ascochyta blight from G24404 in this location of the genome would be promising since no resistance genes for this disease have ever been tagged in common beans but further inoculated and controlled testing would be needed to confirm this. Meanwhile, several of the QTL identified in this study were consistent with QTL from previous studies in common bean. The association of plant height with the indeterminacy gene on linkage group b01 and the linkage of the indeterminacy gene with loci for flowering time and photoperiod response is well established (Koinange et al. 1996). Flowering time QTL that are associated with yield QTL have also been identified on linkage group b09 in a cultivated × cultivated, Mesoamerican recombinant inbred line population (Tar'an et al. 2002). The presence of seed weight QTL on b07 and b11 has also been observed before in reports of intra- and inter-gene pool crosses (Tar'an et al. 2002; Park et al. 2000) and the cultivated × wild combination studied by Koinange et al. (1996).

Utility of advanced backcross method in common beans

The results of this study suggest that wild beans can be a source of genes for higher yield and yield component

traits in cultivated beans and that the advanced backcross strategy can be successful at transferring these genes into cultivars with commercial seed types. Similar results were found with rice and tomato where the same strategy was applied (Tanksley and McCouch 1997). Advanced backcrossing was advantageous for use with wild common beans because no fertility barriers were observed between the wild and cultivated gene pools, the derived lines themselves were close to commercial type, and new diversity for plant vigor and tolerance mechanisms was incorporated into the cultivated background used in this study. The advanced backcross method also had advantages as a first step toward near isogenic line development and QTL gene cloning and in the production of selected breeding lines useful for our genetic improvement program as first discussed by Tanksley and Nelson (1996). QTL were detected even in regions of low introgression indicating that the advanced backcross method is robust for QTL detection in common bean as has been previously shown by authors in other crop. However, one negative aspect of using wild beans for advanced backcrossing was precisely the reduced possibility of introgression from certain regions of the genome especially around domestication genes. In addition, several negative QTL from the wild accession remained in the population where selection for return to recurrent parent genome was not effective. Positive and negative alleles for adaptation and productivity from wild beans were also identified in spontaneous wild by cultivated hybrids that exist in the Andes (Beebe et al. 1997) and in additional advanced backcross populations from CIAT (Buendia et al. 2003; Beaver et al. 2003) although this is the first study to link positive alleles for yield and yield component traits from wild accessions with genetic markers. Despite these results, trait compensation between smaller seed size or later flowering and higher yield remains an important issue that was observed before in wide crosses with wild accessions (Motto et al. 1978; Singh et al. 1995) and which must be considered when further applying advanced backcross breeding to common beans. However, overall it appears that Andean gene pool varieties may be likely to benefit from advanced backcrossing with the wild gene pool more than Mesoamerican gene pool varieties because of their relatively lower level of genetic diversity (Beebe et al. 2001; Buendia et al. 2003).

Uniqueness of the wild accession

Part of the success of this project may have relied on the selection of unique wild progenitor from Colombia that was relatively distant genetically from cultivated beans in general and from Andean cultivated beans in particular. The pattern of polymorphism presented by G24404 compared to the recurrent parent Cerinza allowed for efficient development of a genetic map for the advanced backcross population. In a study of AFLP diversity, G24404 was genetically distinct from both Mesoamerican and Andean wild common bean genotypes and presented infrequent AFLP alleles found only in Northern Andean wild common beans and not in either of the major gene pools (Tohme et al. 1996). G24404 and the Northern Andean wild gene pool were also singled out by Debouck et al. (1993) and by Beebe et al. (2001) who found it to be distinct from other Andean wild beans. This genotype can therefore be considered to represent a unique gene pool of wild *P. vulgaris*, unrelated to the more widely distributed Andean and Mesoamerican gene pools that were widely used in domestication of the crop and as such is a valuable genetic resource for the improvement of cultivars through breeding techniques such as the advanced backcross method presented here.

Acknowledgements We are grateful to Didier Peltier for providing selected microsatellite markers, to Alcides Hincapie for hybridizations between wild and cultivated parents, to Agobardo Hoyos for help with data analysis, to Yercil Viera for field trial management, to Gustavo Vallejo, Daniel Debouck and César Martinez for helpful reviewing of the manuscript or drafts and P. Zamorano for formatting.

References

Afanador LK, Hadley SD (1993) Adoption of a mini-prep DNA extraction method for RAPD marker analysis in common bean. Bean Improv Coop 35:10–11

Anthony JL, Vorder RA, Hall TC (1990) Nucleotide sequence of an alpha-phaseolin gene from *Phaseolus vulgaris*. Nucleic Acids Res 18:3396

Bassett MJ (1991) A revised linkage map of common bean. HortScience 26:834–836

Basten C, Weir BS, Zeng Z-B (2001) Windows QTL Cartographer v 1.21. Software. Department of Statistics, North Carolina State University, USA

Beaver JS (1999) Improvement of large-seeded race Nueva Granada cultivars. In: Singh SP (ed) Common bean improvement in the twenty-first century. Kluwer, Dordrecht

Beaver JS, Rosas JCM, Myers J, Acosta J, Kelly JD, Nchimbi-Msolla S, Misangu R, Bokosi J, Temple S, Aranud-Santana E, Coyne DP (2003) Contributions of the bean/cowpea CRSP to cultivar and germplasm development in common bean. Field Crops Res 82:87–102

Becerra-Velazquez L, Gepts P (1994) RFLP diversity of common bean (*Phaseolus vulgaris* L.) in its centres of origin. Genome 37:256–263

Beebe S, Toro O, González AV, Chacón MI, Debouck DG (1997) Wild-weed-crop complexes of common bean (*Phaseolus vulgaris* L., Fabaceae) in the Andes of Peru and Colombia, and their implications for conservation and breeding. Genet Resour Crop Evol 44:73–91

Beebe SE, Réngifo J, Gaitán-Solis E, Duque MC, Tohme J (2001) Diversity and origin of Andean landraces of common bean. Crop Sci 41:854–862

Blair MW, Pedraza F, Buendía HF, Gaitán-Solís E, Beebe SE, Gepts P, Tohme J (2003) Development of a genome-wide anchored microsatellite map for common bean (*Phaseolus vulgaris* L.). Theor Appl Genet 107:1362–1374

Broughton WJ, Hernandez G, Blair MW, Beebe SE, Gepts P, Vanderleyden J (2003) Beans (*Phaseolus* spp.)—model food legumes. Plant Soil 252:55–128

Buendia HF, Beebe SE, Blair MW, Teran H, Pedraza F (2003) Identificación de marcadores moleculares asociados a genes de rendimiento en una población RC2F3.7 de frijol común *Phaseolus vulgaris* L. (DOR390 × G19892). Fitotecnia Colombiana 3:57–64

Cattan-Toupance I, Michalakis Y, Neema C (1998) Genetic structure of wild bean populations in their South-Andean centre of origin. Theor Appl Genet 96:844–851

Chacón MI, Pickersgill SB, Debouck DG (2005) Domestication patterns in common bean (*Phaseolus vulgaris* L.) and the origin of the Mesoamerican and Andean cultivated races. Theor Appl Genet 110:432–444

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

Coyne DP (1967) Photoperiodism: inheritance and linkage studies in *Phaseolus vulgaris*. J Hered 58:313–314

Coyne DP (1970) Genetic control of a photoperiod-temperature response for time of flowering in beans (*Phaseolus vulgaris* L.). Crop Sci 10:246–248

Debouck DG, Smartt J (1995) Beans: *Phaseolus* spp. (Leguminosae-papilionoideae). In: Smartt J, Simmonds NW (eds) Evolution of crop plants. 2nd edn. Longman, London, pp 297–294

Debouck DG, Tohme J (1989) Implications for bean breeders of studies on the origins of common beans, *Phaseolus vulgaris* L. In: Beebe S (ed) Current topics in breeding of common bean, vol 47. Bean program, (CIAT), Cali, Colombia, pp 3–42

Debouck DG, Toro O, Paredes M, Johnson WC, Gepts P (1993) Genetic diversity and ecological distribution of *Phaseolus vulgaris* (Fabaceae) in northwestern South America. Econ Bot 47:408–423

Freyre R, Ríos R, Guzmán L, Debouck D, Gepts P (1996) Ecogeographic distribution of *Phaseolus* spp. (Fabaceae) in Bolivia. Econ Bot 50:195–215

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

Gaitán-Solís E, Duque MC, Edwards KJ, Tohme J (2002) Microsatellite repeats in common bean (*Phaseolus vulgaris*): isolation, characterization, and cross-species amplification in *Phaseolus* ssp. Crop Sci 42:2128–2136

Gepts P (1988) Phaseolin as an evolutionary marker. In: Gepts P (ed) Genetic resources of *Phaseolus* beans; their maintenance, domestication, evolution, and utilization. Current plant science and biotechnology in agriculture. Kluwer, Dordrecht, 624 pp

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

Gepts P (2002) A comparison between crop domestication, classical plant breeding, and genetic engineering. Crop Sci 42:1780–1790

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

Gepts P, Debouck D (1991) Origin, domestication, and evolution of the common bean (*Phaseolus vulgaris* L.). In: Schoonhoven AV (ed) Common beans: research for crop improvement. CIAT, Cali, pp 7–53

Gu W, Zhu JQ, Wallace DH, Singh SP, Weeden NF (1998) Analysis of genes controlling photoperiod sensitivity in common bean using DNA markers. Euphytica 102:125–132

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 125:69–72

Kami J, Becerra Velázquez V, Debouck DG, Gepts P (1995) Identification of presumed ancestral DNA sequences of phaseolin in *Phaseolus vulgaris*. Proc Natl Acad Sci USA 92:1101–1104

Kelly JD, Schneider KA, Kolkman JM (1999) Breeding to improve yield. In: Singh SP (ed) Developments in plant breeding. Common bean improvement in the twenty-first century. Kluwer, Dordrecht

Kelly JD, Gepts P, Miklas PN, Coyne DP (2003) Tagging and mapping of genes and QTL and molecular-marker assisted selection for traits of economic importance in bean and cowpea. Field Crops Res 82:135–154

Khairallah M, Sears B, Adams M (1992) Mitochondrial restriction fragment length polymorphisms in wild *Phaseolus vulgaris* L.: insights on the domestication of the common bean. Theor Appl Genet 84:915–922

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

Koenig RL, Singh SP, Gepts P (1990) Novel phaseolin types in wild and cultivated common bean (*Phaseolus vulgaris*, Fabaceae). Econ Bot 44:50–60

Koinange EMK, Singh SP, Gepts P (1996) Genetic control of the domestication syndrome in common bean. Crop Sci 36:1037–1045

Kornegay J, Cardona C, Posso CE (1993) Inheritance of resistance to Mexican bean weevil in common bean, determined by bioassay and biochemical tests. Crop Sci 33:589–984

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

López CE, Acosta IF, Jara C, Pedraza F, Gaitán-Solís E, Gallego G, Beebe S, Tohme J (2003) Identifying resistance gene analogs associated with resistances to different pathogens in common bean. Phytopathology 93:88–95

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, 980 pp

McClean PE, Lee RK, Otto C, Gepts P, Bassett MJ (2002) Molecular and phenotypic mapping of genes controlling seed coat pattern and color in common bean (*Phaseolus vulgaris* L.). J Hered 93:148–152

Métais I, Hamon B, Jalouzot R, Peltier D (2002) Structure and level of genetic diversity in various bean types evidenced with microsatellite markers isolated from a genomic enriched library. Theor Appl Genet 104:1346–1352

Miklas PN, Coyne DP, Grafton KF, Mutlu N, Reiser J, Lindgren DT, Singh SP (2003) A major QTL for common bacterial blight resistance derives from the common bean great northern landrace cultivar Montana No. 5. Euphytica 131:137–146

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

Nodari RO, Tsai SM, Gilbertson 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

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

Singh SP (2001) Broadening the genetic base of common bean cultivars: a review. Crop Sci 41:659–1675

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

Singh S, Nodari R, Gepts P (1991b) Genetic diversity in cultivated common bean. II. Marker-based analysis of morphological and agronomic traits. Crop Sci 31:19–23

Singh SP, Molina A, Gepts P (1995) Potential of wild common bean for seed yield improvement of cultivars in the tropics. Can J Plant Sci 75:807–812

Sullivan JG, Bliss FA (1983) Expression of enhanced seed protein content in inbred backcross lines of common bean. J Am Soc Hort Sci 108:787–791

Tanksley SD, McCouch SR (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. Science 277:1063–1066

Tanksley SD, Nelson JC (1996) Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTL from germplasm into elite breeding lines. Theor Appl Genet 92:191–203

Tar'an B, Michaels TE, Pauls KP (2002) Genetic mapping of agronomic traits in common bean. Crop Sci 42:544–556

Tohme J, Gonzlez DO, Beebe S, Duque MC (1996) AFLP analysis of gene pools of a wild bean core collection. Crop Sci 36:1375–1384

Vallejos C, Sakiyama NS, Chase CD (1992) A molecular marker based linkage map of *Phaseolus vulgaris* L. Genetics 131:733–740

Voysest O (2000) Mejoramiento genético del frijol (*Phaseolus vulgaris* L.): Legado de variedades de América Latina 1930–1999. CIAT, Cali

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

Yu K, Park SJ, Poysa V, Gepts P (2000) Integration of simple sequence repeat (SSR) markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.). J Hered 91:429–434