

Use of the advanced backcross-QTL method to transfer seed mineral accumulation nutrition traits from wild to Andean cultivated common beans

Matthew W. Blair · Paulo Izquierdo

Received: 9 February 2012 / Accepted: 27 April 2012 / Published online: 21 June 2012
© Springer-Verlag 2012

Abstract Iron deficiency anemia and zinc deficiency are major health concerns across the world and can be addressed by biofortification breeding of higher mineral concentration in staple crops, such as common bean. Wild common beans have for the most part had higher average seed mineral concentration than cultivars of this species but have small un-commercial seeds. A logical approach for the transfer of the seed mineral trait from wild beans to cultivated beans is through the advanced backcross breeding approach. The goal of this study was to analyze a population of 138 BC₂F_{3,5} introgression lines derived from the very high iron wild genotype G10022 backcrossed into the genetic background of the commercial-type variety ‘Cerinza’, a large-red seeded bush bean cultivar of the Andean gene pool. In addition to measuring seed mineral accumulation traits and the quantitative trait loci (QTL) controlling these traits we were interested in simultaneously testing the adaptation of the introgression lines in two replicated yield trials. We found the cross to have high polymorphism and constructed an anchored microsatellite map for the population that was 1,554-cM long and covered all 11 linkage groups of the common bean genome.

Through composite interval mapping (CIM) and single point analysis (SPA), we identified associations of markers and mineral traits on b01, b06, b07, b08, b10 and b11 for seed iron concentration, and markers on b01, b04 and b10 for seed zinc concentration. The b07 and b08 QTL aligned with previous QTL for iron concentration. A large number of QTL were found for seed weight (9 with CIM and 36 with SPA analysis) and correlations between seed size and mineral content affected the identification of iron and zinc contents’ QTL on many linkage groups. Segregation distortion around domestication genes made some areas difficult to introgress. However, in conclusion, the advanced backcross program produced some introgression lines with high mineral accumulation traits using a wild donor parent.

Introduction

Nutrition traits are important breeding goals for most staple crops and have been incorporated into international breeding programs through the concepts of biofortification (Pfeiffer and McClaferty 2007). This project addresses “hidden hunger” caused by micronutrient deficiencies through breeding for higher mineral and vitamin status of cereals, pulses and tuber crops. Iron deficiency anemia (IDA) is especially a common micronutrient deficiency affecting over 2 billion people around the world which is caused by low iron intake, especially in reproductive age women and developing adolescents. IDA is difficult to address through iron supplementation or processed foods; therefore, several attempts are being made to breed high iron accumulation into staples such as rice, maize, wheat and legumes.

Common bean (*Phaseolus vulgaris* L.) is the most widely grown grain legume for direct human consumption

Communicated by C. Gebhardt.

M. W. Blair
Universidad Nacional de Colombia,
Km 12 via Chapinero, Palmira, Valle, Colombia

M. W. Blair (✉)
Department of Plant Breeding and Genetics,
Cornell University, 242 Emerson Hall, Ithaca, NY 14853, USA
e-mail: mwbeans@gmail.com; mwb1@cornell.edu

P. Izquierdo
CENICAÑA, Centro Nacional de Investigación en Caña,
Candelaria, Valle, Colombia

and is highly preferred in many parts of Africa and Latin America, as well as in southern Europe (Broughton et al. 2003). This pulse is part of the healthy diet of the Mediterranean basin and is of growing importance in the United States where consumption has been increasing due to greater interest in “ethnic” and healthy foods. Consumption rates for the crop can be as high as 40–60 kg/year in regions where meat is scarce such as in the Great Lakes of Africa or in rural northeast Brasil, Central America and Mexico. In various countries of eastern Africa, common beans make up the most important source of dietary protein (Wortmann et al. 1998).

Archeological evidence shows that common beans were domesticated in the New World over 5,000–8,000 years ago and that domestication took place in both the Andes of South America and Mesoamerica (Singh et al. 1995). These domestications in the northern and southern ranges of the wild common bean gave rise to the Andean and Mesoamerican gene pools of cultivated accessions, respectively. Wild *P. vulgaris* accessions are considered a primary gene pool for common bean improvement because they are easily crossed with cultivars and yet still have traits that were not incorporated into the cultivated gene pool (Singh et al. 1995). Furthermore, the genetic diversity of wild beans is thought to be broader than cultivated beans (Singh 2001).

Nutritional breeding of common bean started in the 1980s with selection for seed protein types and percentages (Sullivan and Bliss 1983; Delaney and Bliss 1991). At that time, amino acid profiles were suggested as the main element to study. Starting in the late 1990s, mineral concentration was added as a component of nutritional breeding and an extensive germplasm collection was screened (Islam et al. 2002). High seed mineral accumulation was found in a set of landraces with high seed iron and zinc concentrations and these were used to create mapping populations (see Blair et al. 2009a). Apart from this, screening of wild and weedy beans identified genotypes from Mexico such as G10022 with high iron concentration.

Inheritance studies for iron and zinc accumulation have been undertaken so far in six populations of common bean recombinant inbred lines. These quantitative trait loci (QTL) studies have shown that the inheritance is quantitative and that a range of different loci in the different populations control seed iron and zinc accumulation. Among the QTL populations evaluated so far, one has been from an inter gene pool cross (Blair et al. 2009b), two from Andean \times Andean gene pool crosses (Blair et al. 2011a; Cichy et al. 2009) and one from a Mesoamerican \times Mesoamerican cross (Blair et al. 2010b). All of these QTL analyses are cross-comparable by way of microsatellite marker-anchored genetic maps. Another recombinant inbred line population was developed in navy beans to tag a

major gene for zinc deficiency tolerance and accumulation of seed zinc but the associated genetic map had only partial genome coverage (Gelin et al. 2007). To date, no other types of common bean populations (backcross, F2 or near isogenic lines) have been studied for nutrition traits.

Apart from the cultivar populations described above, one wild \times cultivated recombinant inbred line population was studied for nutrient and anti-nutrient traits, including iron and tannin accumulation, respectively (Guzmán-Maldonado et al. 2003). The genetic map for the latter population is not cross-comparable with the anchored maps due to a lack of single-copy markers, but the population is of interest because it used a wild source of high iron. Furthermore, the high iron trait seems to be prevalent in wild accessions of common bean (this laboratory, unpublished data).

In architectural terms, wild beans are spindly, type IV indeterminate growth habit and therefore, may be more useful for climbing bean improvement than for bush bean breeding (Blair et al. 2006b). Other physiological differences based on partitioning and leaf, seed and other organ sizes make wild beans difficult to work with. Finally, a suite of domestication traits such as non-shattering pods, photoperiod insensitivity and light rather than dark colored seeds are associated with specific genes (Koinange et al. 1996) which upon selection for cultivated type genotypes cause selective sweeps and segregation distortion in populations derived from wild \times cultivated bean crosses (Blair et al. 2006b).

The difficulty of using wild beans indicates that a method is needed to capture interesting traits and genes from the wild accessions without linkage drag. Such a method is through backcrossing where a simple trait is to be transferred or inbred backcrossing where a complex trait is to be transferred (Sullivan and Bliss 1983). A modification of the inbred backcross method is the advanced backcross (AB) method of applying QTL analysis simultaneously to plant-to-plant crosses to evaluate and assist in gene transfer (Tanksley and Nelson 1996). One important advantage of the AB method is that favorable wild alleles can be transferred from unadapted germplasm into elite breeding lines while avoiding the epistatic effects of deleterious background genes found in the wild (Tanksley and McCouch 1997).

Combined with QTL analysis, an AB population can be used to identify positive alleles from the wild which can be transferred through further crosses by marker-assisted selection (Tanksley and Nelson 1996). An additional advantage of the AB method is that it produces chromosome segment substitution line libraries (Ghesquiere et al. 1997). The AB-QTL method has been widely exploited in barley, tomato and rice (other inbreeding crops) to move genes and traits of interest from wild sources to cultivated

backgrounds (Ando et al. 2008; Eshed and Zamir 1995; Fulton et al. 1997; Lin et al. 1998) but has only been used once to move agronomic traits from the wild into cultivated common beans (Blair et al. 2006b). To our knowledge, the AB-QTL method has never before been used for the transfer of a micronutrient trait in any crop although wild relatives are considered sources of high mineral traits in several cereals as well as in common bean (Graham et al. 1999; Guzmán-Maldonado et al. 2004).

The objectives of this study, therefore, were (1) to genotype an AB population for a high seed mineral wild bean in an Andean genepool-cultivated common bean background, (2) to evaluate the resulting introgression lines for seed mineral (iron and zinc) concentrations and (3) to determine the QTL controlling seed mineral accumulation in the population. The study here is among the first to attempt to obtain higher seed mineral concentration from wild beans to improve cultivated types.

Materials and methods

Plant materials

An advanced backcross population was developed by crossing the cultivated recurrent parent Cerinza with the wild donor parent G10022, a source of high seed mineral concentration. A description of the recurrent parent Cerinza, a released cultivar from Colombia with large red seeds from the Andean genepool, is found in Blair et al. (2006a, b). Meanwhile, the wild donor parent G10022 from Mexico was discovered in a screening program for nutritional traits in wild beans carried out by Islam et al. (2002). Cerinza was chosen because it had been used to create another AB population described in Blair et al. (2006b) and was an appropriate representative of the Andean genepool. Like other Andean large red beans, Cerinza has a good baseline of iron concentration (55–72 ppm) but could be improved for this trait (Astudillo and Blair 2008). G10022 was selected as the donor parent because it was exceptionally high in iron concentration among wild accessions and cultivated landraces showing 108 ppm iron concentration compared to the average concentration of 55 ppm for common bean. Zinc concentration was less contrasting in previous studies of the two parents with G10022 having 38 ppm and Cerinza having 25–27 ppm, although due to phenological differences (late flowering of the wild accession) the genotypes were never grown simultaneously in the same growing environment to compare mineral accumulation.

Like most wild beans, G10022 is a type IV climbing bean and produces very small (under 15 g/100 seeds) seed that is mottled light gray, dark gray and black in color and

square-ended or flattened in shape. It flowers later than cultivated beans at 95–110 days after planting when sown in September in Colombia and matures irregularly after 140 days after planting. G10022 has a darker hilum than the surrounding seed coat and was collected from a wild bean population found in Durango, Mexico at a site at 1,829 m above the sea level (masl) elevation. This wild accession has the “S” phaseolin seed protein pattern which makes it typical of Mesoamerican genotypes and part of the opposite genepool to Cerinza. G10022 is conserved in the Genetic Resources Unit—FAO collection found in Cali, Colombia.

For population development an initial backcross was made to the cultivated parent using Cerinza and the F_1 from the initial simple cross as described in Blair et al. (2006b). As many pollinations were made as possible so as to obtain over 180 BC_1F_1 seeds. The resulting seed was planted out in an insect-proof greenhouse in Popayán, Colombia (18 °C average temperature) and crossed plant-to-plant with the recurrent parent to again develop over 180 BC_2F_1 seeds from different individual plants.

The resulting seed was planted out to produce separate BC_2F_1 plants for single-plant harvests. The $BC_2F_{1,2}$ families were then planted out in a field environment to produce sufficient seed for single-plant selections (SPS) in the BC_2F_2 generation. Seed of each SPS was then multiplied through two more generations to have sufficient seed for replicated experiments at two locations as described below and a total of 138 $BC_2F_{2,5}$ lines met these criteria and were used in the agronomic trials.

The $BC_2F_{2,5}$ 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_2 generation of inbreeding was favored for the detection of additive genes. Generation advance to the $F_{2.5}$ 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 about 30 seeds harvested per seed planted due to its large seed size. Most segregants were large-seeded given the use of the BC_2 generation which brought the genotype toward the cultivated recurrent parental phenotype. The only selection was against very small or misshapen seed but variability in seed size was allowed.

Agronomic trials and mineral analysis

The population was grown at two locations with contrasting soil and temperature conditions. The first location used was Palmira, Colombia (996 masl, average yearly temperature of 24 °C, average relative humidity 70 % and average yearly rainfall 950 mm) with neutral pH, predominantly clay soils (pH 7.2, Vertisol). Meanwhile, the

second location was Darién, Colombia (3°54'N, 76°30'W, 1,485 masl, average yearly temperature 20 °C, average relative humidity 80 % and average yearly rainfall 1,288 mm) with moderately acid, loam soils (pH 5.6, Andisol).

Both the locations had approximately 12-h photoperiod at flowering and fertilizations with boron and zinc chelate micronutrients in three bi-weekly foliar sprays were used for plant establishment. Seed protection was provided by Benlate-Benomyl fungicide dressing. Agronomic management consisted in manual weeding and hilling in Darién versus mechanical tillage in Palmira. Since the soils in Darién are of low fertility, pre-planting application of super-phosphate fertilizer was used there, but not in Palmira where soils are of high fertility. Insecticide control of whitefly and leafhoppers was required in Palmira but not in Darién. For both the locations, plots consisted of two rows that 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 randomized complete block design with three repetitions.

Trait evaluation and data analysis

Mineral analysis was the primary goal of the study and therefore seed samples from each replicate were analyzed for iron and zinc concentration. In addition, average seed weight (g) for 100 seeds and yield on a per plot basis were evaluated for each treatment. Yields were also calculated on a kg per hectare (ha) basis based on the harvested plant number as a co-variable. In all harvests, the last two plants in each row were not considered to avoid border effects. Days to flowering and to maturity were evaluated as phenological traits but were relatively uniform and were not used further.

Mineral analysis was carried out for all repetitions and locations by atomic absorption spectroscopy (AAS) while a mixed sample from the three replications of the Palmira harvest was also analyzed by inductively coupled plasma (ICP) analysis as described in Blair et al. (2009b). For each repetition, 5 g of seed was washed with sterile double-distilled water and then dried for 24 h at 45 °C in a bench-top oven, before grinding in a modified Retsh mill with 24 sample slots using zirconium grinding balls and Teflon grinding chambers. AAS analysis was carried out with samples digested in 37 % HCl acid that were vaporized and evaluated on a UNICAM969 spectrophotometer at the International Center for Tropical Agriculture analytical laboratory in Colombia. ICP analysis was carried out on optical emission spectroscopy (EOS) equipment at the Waite Center in the University of Adelaide, Australia.

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 be used for subsequent QTL analysis. Population distributions were evaluated for skewing and normality using the same software.

Molecular marker analysis

DNA extraction was carried out with the method of Afanador et al. (1993) based on a bulked tissue sample from 10 seedlings for each line and each parent grown in a greenhouse to avoid pest or pathogen contamination. DNA quality was checked on a 0.8 % agarose TAE gel and quantified with Quantity One v. 4.3 software. The initial DNA was diluted to a 10-ng/ul for use in PCRs.

A parental survey was then carried out with a total of 261 microsatellite markers (Blair et al. 2003, 2006a, 2008, 2009b, c, d; Grisi et al. 2007a, b; Métais et al. 2002; Yu et al. 2000) from which the polymorphic markers were selected to run on the entire population. All the markers were amplified on PTC-200 thermocyclers (MJ research) using standard reagents (12.5 ng of genomic DNA, 2.5 mM of MgCl₂, 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).

The PCR products were then mixed with 5 µl of loading buffer (formamide mixed with 0.4 % w/v bromophenol blue and 0.25 % w/v xylene cyanol) per well and denatured at 94 °C for 1 min before analysis. Microsatellite markers were then run on Owl T-Rex S3S (ThermoFisher Scientific Inc., USA) sequencers in 4 % polyacrylamide (29:1 acrylamide: bisacrylamide) gels which were silver stained upon completion of the runs. 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.

Genetic mapping and QTL analysis

Linkage analysis (using the 'group' and 'map' commands) was performed with Mapmaker v. 3.0 software (Lander et al. 1987). Linkage group designations for each marker were checked against known microsatellite marker positions on other common bean genetic maps constructed by Blair et al. (2003, 2008) and Grisi et al. (2007a, b). Un-mapped markers were placed using the 'try' command and markers were assigned with a minimum LOD value of 3.0. Map distances are reported in centiMorgans (cM) estimated with the Kosambi mapping function. Chi-square tests were used to determine segregation distortion from the ratio of Cerinza and G10022 alleles expected for the BC₂F_{2.5} generation.

QTL were identified using two software programs: qGene v. 3.07A (Nelson 1997) for single point analysis (SPA) and QTL Cartographer v. 2.5 (Basten et al., 2001) for composite interval mapping analysis (CIM). In both the cases we used the same newly constructed genetic map and the phenotypic data defined above. In the CIM analysis, we performed the analysis every 1 cM (walkspeed) with a window size of 10 cM and using 10 background markers in a forward–backward stepwise multiple linear regression model. In terms of population type, the B12 genetic model was assumed for the CIM analysis. Meanwhile the similar BC₂F₁ model was assumed for the SPA analysis. In both cases, the homozygous donor parent allele class was combined with the heterozygous genotypic class.

In terms of probability threshold, for the SPA analysis we used a threshold of 0.001 % while for the CIM analysis we determined a LOD threshold with a 1,000 permutation test as recommended by Churchill and Doerge (1994). 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 for micronutrients were named as using the mineral name and a two number code derived from the linkage group and the number of the QTL identified on that linkage group, separated by a period. Content-based QTL names for seed minerals were distinguished with an underscore designation of *Fe_cont* or *Zn_cont* as in Blair et al. (2010b).

In addition, introgression analysis was carried out with the software program CSSL Finder v. 0.5.3. Here, we used the ‘compute stats’ command to determine the number of donor parent chromosomal segments and their location. Finally, genetic maps and QTL locations were drawn with MapChart v. 1.0 software program (Voorrips 2002).

Results

Seed mineral and yield trait variability

Seed yield and mineral concentration trait variability was ample for the genotypes in the advanced backcross population showing the quantitative nature of inheritance for these traits. Analyses of variance showed that environmental effects and genotypic differences were significant for all traits. Genotype × environment (G×E) effects were not significant for zinc concentration and seed weight, while seed yield and iron concentration were influenced by G×E effects as well as location effects.

Cerinza had values for seed yield and mineral concentrations close to the average of the BC₂F_{2.5} lines in both the

locations but was higher yielding in Darién (1,520 kg/ha) than Palmira (1,338 kg/ha). Cerinza was also high in iron (92 ppm) and zinc (35 ppm) concentration in Palmira but near or below the average of the lines in Darién (70 ppm and 29 ppm for iron and zinc, respectively). Cerinza was higher in iron and zinc compared to previous values of 60 ppm iron and 25 ppm zinc in a previous study (Astudillo and Blair 2008) and this was unexpected. For seed weight, Cerinza (53 g/100 seeds) was close to the top end of the population trait range in both locations with the average of the BC₂F_{2.5} lines higher in the Darién (44 g) than in the Palmira (39.5 g).

The maximum performing BC₂F_{2.5} line for yield was also higher in Darién (2,159 kg/ha) than in Palmira (1,840 kg/ha); however, the maximum seed weight in both locations was similar (56.6 and 56.0 g for Darién and Palmira, respectively). In each location, the population produced lines that were superior in yield by up to 37–42 % over the recurrent parent. The worst performing lines had 92 and 49 % of the yields of Cerinza in Darién and Palmira, respectively. Overall the average yields of the lines were higher in Darién than in Palmira by about 10 %. Seed weight increases over Cerinza were small with only 5.3 and 5.7 % increases observed in Darién and Palmira, respectively. A range of lines had seed weights significantly under Cerinza as would be expected in a large population of advanced backcross lines.

Overall the population averages in each location tended to be skewed toward larger seed size compared to the donor parent which has very small seed size (≥12 g/100 seeds). This was because segregation in the backcross population was toward the large grain typical of Cerinza as recurrent parent. In addition, seed size was higher in Darién than in Palmira due to the cooler growing condition there which are more conducive to seed filling and this was observed in the differences for the average of all the advanced backcross lines even more than for Cerinza which was stable across the two locations.

Iron concentration of the most mineral dense BC₂F_{2.5} line was 99 ppm in Palmira which represented a small increase over the recurrent parent (8 % increase). However in Darién, differences were more significant (19 % increase), although the highest seed iron line was only 92 ppm. In terms of zinc concentration the most mineral dense line was 38 ppm in Darién and 39 ppm in Palmira, representing 44 and 12 % increases, respectively, over the recurrent parent. Low mineral lines ranged down to 58–56 ppm for iron concentration in Darién and Palmira, respectively, and down to 23–27 ppm for zinc concentration in the two sites. In terms of the average of all the lines, iron concentration was generally higher in Palmira than in Darién possibly due to soil differences in the first site with neutral pH and higher soil iron concentration (3.5 ppm)

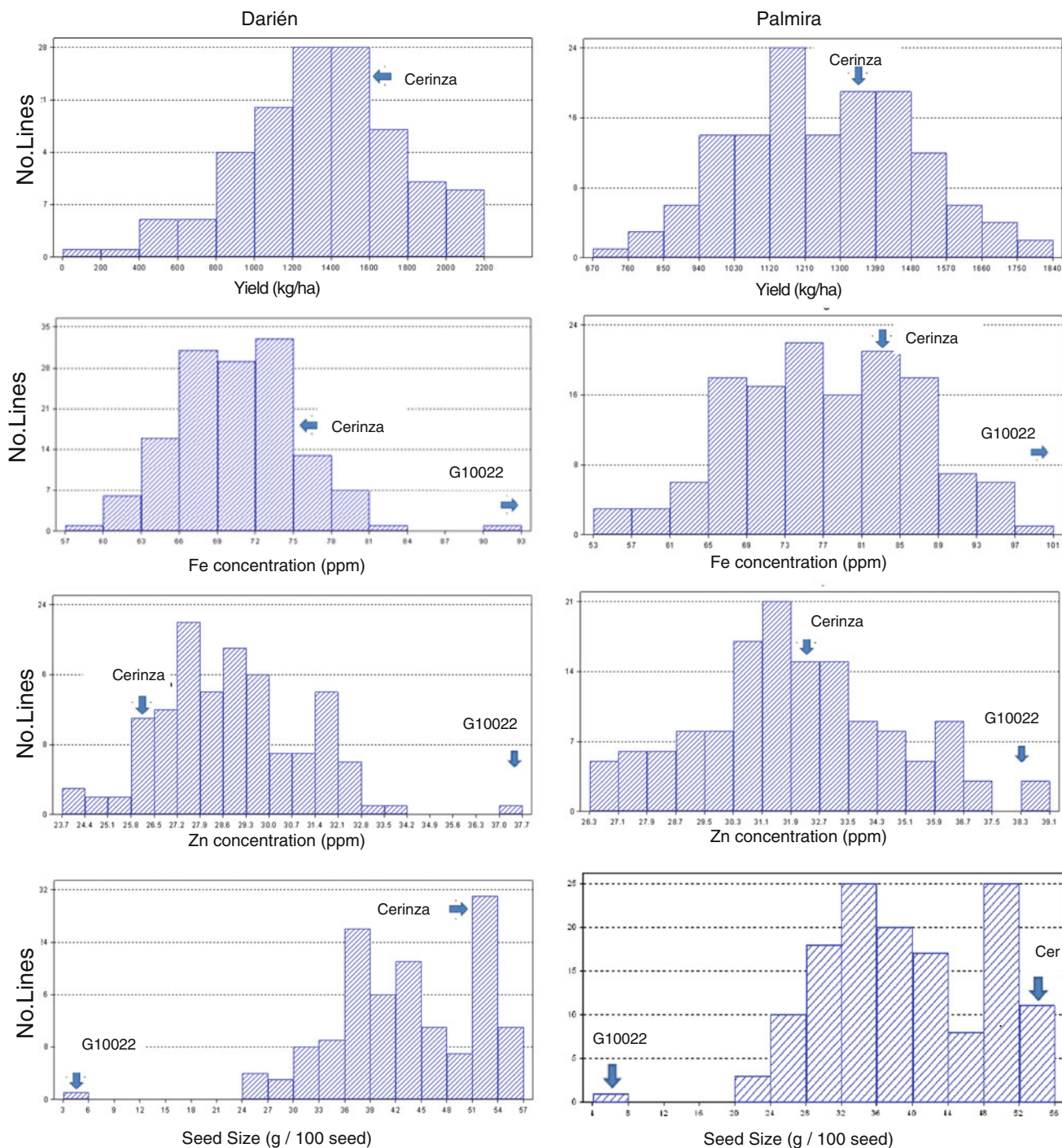


Fig. 1 Population frequency distributions for seed yield (kg/ha), seed mineral concentrations (ppm) and seed weight (g/100 seeds) in the (Cerinza \times (Cerinza \times (Cerinza \times G10022))) advanced backcross

versus the second site which has acid pH soils and lower soil iron concentration (1.7 ppm). The opposite was the case for zinc concentration where soil zinc deficiency (0.9 ppm soil concentration) is characteristic of Darién. Meanwhile, zinc concentrations were generally lower in

population. Iron (Fe) and zinc (Zn) concentrations based on atomic absorption spectrophotometry measurements. *Arrows* indicate phenotypic value of recurrent parent ICA Cerinza

Darién than in Palmira where soil zinc deficiency is less of a problem (2.7 ppm soil concentration).

Population histograms (Fig. 1) confirm the quantitative nature of the iron and zinc concentration traits. Specifically, normal distributions were observed in both locations

and for both minerals. The wide segregation also showed that introgression of phenotypic characteristics from the wild accession into the cultivated background had occurred in the development of the advanced backcross population. Correlation analysis showed that iron and zinc concentrations were correlated with each other in both the locations ($r = 0.51$ for Darién and $r = 0.35$ for Palmira, $P \leq 0.001$). Likewise, seed mineral concentrations were correlated across sites, especially for zinc at the two sites ($r = 0.45$). In comparison, iron at the two sites had a less significant correlation ($r = 0.21$).

Utility of microsatellite marker screening

A total of 261 microsatellite markers from various sources were screened for polymorphisms between the parents of the advanced backcross population (the cultivated recurrent parent ICA Cerinza and the wild donor parent G10022). The polymorphism rate was found to be high both for genomic (63.5 %) and gene-based (72.3 %) microsatellites providing a large number of microsatellites for amplification in the entire mapping population. Overall, 173 microsatellites (66.3 % of the total) were screened across the full set of advanced backcross lines, of which 151 provided complete datasets useful for genetic mapping.

The resulting genetic map constructed for the advanced backcross population had a total length of 1,554.2 cM length distributed in 11 linkage groups corresponding to the chromosome number in common bean ($n = 11$). A total of 143 microsatellite markers were linked to the map and nine were unlinked. The linkage groups ranged in length from 65.3 cM (b05) to 222.6 cM (b03) with 8–20 markers each. The linkage groups with the largest number of markers present were b01 and b02 while the linkage groups with the smallest number of markers were b05 and b06. Overall, the average linkage group length was 141.3 cM with 13 markers each. Average distance between markers on the genetic map was 11.1 cM which made this map ideal for QTL detection.

Previous studies have also found a large number of markers on linkage groups b02, b03 and b04 and small number on other linkage groups. It was surprising to find a large number of polymorphic markers on linkage group b01 and this may be due to the use of a wild parent in the cross. In contrast, in most cultivated \times cultivated crosses this linkage group is fixed for the domestication syndrome (Koinange et al. 1996). Meanwhile, linkage group b04 was surprisingly small with only 12 polymorphic markers and total genetic distance of 84.2 cM indicating that it might be of only partial coverage. A large resistance gene cluster may occupy the lower part of this linkage group and this may not have been detected in our polymorphism survey. All other linkage groups, except b05 with low

coverage as well, had distances of more than 100 cM as would be expected for a full coverage map. The linkage groups, b09, b10 and b11, which in other genetic maps tend to be poorly saturated were well represented (11–14 markers, 91.9–202.9 cM).

In this genetic mapping, the microsatellites were anchored to a linkage groups based on previous mapping of the markers in Blair et al. (2003, 2008) and Grisi et al. (2007a, b). Where markers were linked at low LOD they were tried on other linkage groups, but generally the markers agreed with their expected chromosomal position. A total of nine markers were low LOD or caused large gaps and were not included in the genetic map. Synteny of the marker order with previous maps and with Blair et al. (2006b) was confirmed visually.

We found it very useful in this study to include microsatellite markers from different sources. The mainly GA-based genomic SSR markers mapped by Blair et al. (2003) were found to be interspersed between the CA-based genomic SSR markers from Grisi et al. (2007a, b). Similarly, the ATA-based markers from Métais et al. (2002) and Blair et al. (2008) were found to be well distributed among the genomic markers from the BM series (Blair et al. 2003, 2009c). Gene-based markers had a higher polymorphism rate than expected because they were pre-selected to be the more polymorphic markers from previous publications (Blair et al. 2003, 2009d, 2011b). In addition, polymorphism was high given that the cross was based on Andean and Mesoamerican parents. Inter-genepool crosses have been shown to be high in polymorphism for gene-based markers while intra-genepool crosses are not (Blair et al. 2006a).

Given the high rate of polymorphism, the genetic map was easy to saturate with single-copy SSR markers and reached a total length of over 1,500 cM which is the average size of most genetic maps in common bean, given the number of chromosomes ($n = 11$). Previous genetic maps in common bean have ranged in length from 1,200 to 2,000 cM depending on the degree of saturation and the type of populations, with most recombinant inbred lines having larger map size than F2 populations (for example Blair et al. 2003, 2008; Grisi et al. 2007a, b). Finally, the advanced backcross population and its genetic map was useful for placing new co-dominant SSR markers that had never before been mapped such as BM216, BM218, BM221, BM275, and BM284 on linkage group b01, BM253 on b02 and BM239 on b04. The position of these markers should be confirmed in balanced mapping populations.

Segregation distortion and introgression levels

Distortion in allele segregation from the expected ratio for a BC₂F₂ population was observed for markers in several

Table 1 Summary of marker locations on the genetic map constructed for the (Cerinza × (Cerinza × (Cerinza × G10022))) advanced backcross populations

| Linkage group | Number of SSR | cM distance | | Segregation distortion | | SD per LG |
|---------------|---------------|---------------|--------------|------------------------|------------|--------------|
| | | Total | Average | Cerinza | G10022 | |
| b01 | 20 | 189.8 | 9.5 | 6 | 0 | 30.0 |
| b02 | 20 | 189.5 | 9.5 | 8 | 1 (BM164) | 45.0 |
| b03 | 17 | 226.6 | 14.5 | 7 | 1 (BM181) | 47.1 |
| b04 | 12 | 84.2 | 7.0 | 4 | 0 | 30.0 |
| b05 | 8 | 65.3 | 8.2 | 3 | 3 | 75.0 |
| b06 | 8 | 117.2 | 14.7 | 2 | 0 | 25.0 |
| b07 | 11 | 126.8 | 11.5 | 3 | 2 | 45.5 |
| b08 | 10 | 106.8 | 10.7 | 5 | 1 (ATA289) | 60.0 |
| b09 | 14 | 202.9 | 14.5 | 5 | 0 | 35.7 |
| b10 | 11 | 153.2 | 13.9 | 3 | 0 | 27.3 |
| b11 | 12 | 91.9 | 7.7 | 2 | 0 | 16.7 |
| Total | 143 | 1554.2 | 11.06 | 48 | 8 | 39.1 |

Markers with skewed segregation distortion toward G10022 are indicated within the column where only one marker is skewed. Other skewed markers toward the donor parent include BMd20, BMd53 and BM155 on linkage group b05, BM185 and PV35 on linkage group b07

regions of the genome and on specific linkage groups of the newly constructed genetic map (Table 1). A total of over one-third (39.1 %) of the mapped markers were distorted, while nearly two-thirds (60.1 %) were not distorted. Segregation distortion was most notable on linkage group b02 where 9 markers out of 20 (45 %) were skewed in allele frequency, on linkage group b05 where 6 out of 8 markers (75 %) were skewed and on linkage group b08 where 60 % of the markers were skewed. Another linkage group with high number of skewed markers was linkage group b03 with 8 out of 29 markers (47 %) distorted. Some segregation distortion was found on all the additional linkage groups as well. In the majority of cases, skewed allele frequency was toward the recurrent parent alleles. For example, distortion toward the recurrent parent was observed on linkage groups b01, b02, b03, b04, b08, b09 and b10. Meanwhile, distortion toward the donor parent alleles was limited to one region of three markers on linkage group b05 and two markers on linkage group b07. On both of these linkage groups other segments were skewed toward the recurrent parent alleles. In general, the pattern of segregation distortion showed that introgression of wild parent alleles was difficult and selected against during the process of advanced (a.k.a. inbred) backcross selection. However, on certain linkage groups introgression occurred normally and this was reflected in a graphical analysis of the advanced backcross lines discussed below.

Introgression line analysis was carried out with chromosome segment substitution line (CSSL) finder and is shown in Fig. 2. In this graph, the introgression lines (horizontal bars aligned to the left axis) are ordered based on the most non-overlapping chromosomal segments

(intervals as columns aligned to the top axis based on linkage group position) across the entire genome. The segments that are fixed for the donor parent (G10022) are shown in black. Other colors represent segments fixed for the recurrent parent (Cerinza) alleles shown in light gray background color. Meanwhile, those segments still segregating in the BC₂F_{2.5} generation are shown in dark blue. Thanks to the large number of introgression lines analyzed, no region was completely missing an introgressed segment from the donor parent. The high rate of return to the recurrent parent genotype is observable from the predominance of Cerinza segments in the graph.

Quantitative trait loci identified

In the single-point analysis, a total of 10-positive markers were detected for yield QTL on linkage groups b01, b02, b03, b04 and b09 (Table 2). All of these associations were with the positive alleles from the recurrent parent Cerinza. Iron concentration was associated with 5 markers on linkage groups b01, b10 and b11 with the QTL detected in Darién associated with G10022, the wild donor parent, and the QTL detected in Palmira associated with Cerinza. Meanwhile, zinc concentration QTL were associated with eight markers on linkage groups b01, b04, b10 and b11. Four of these associations were found in Darién and four in Palmira, three for AAS and one for ICP. Five of the zinc concentration QTL were associated with positive alleles from G10022 the wild donor parent and three with positive alleles from the recurrent cultivated parent, Cerinza. Finally, a total of 19 markers were associated with net zinc content and these were found on nine linkage groups and

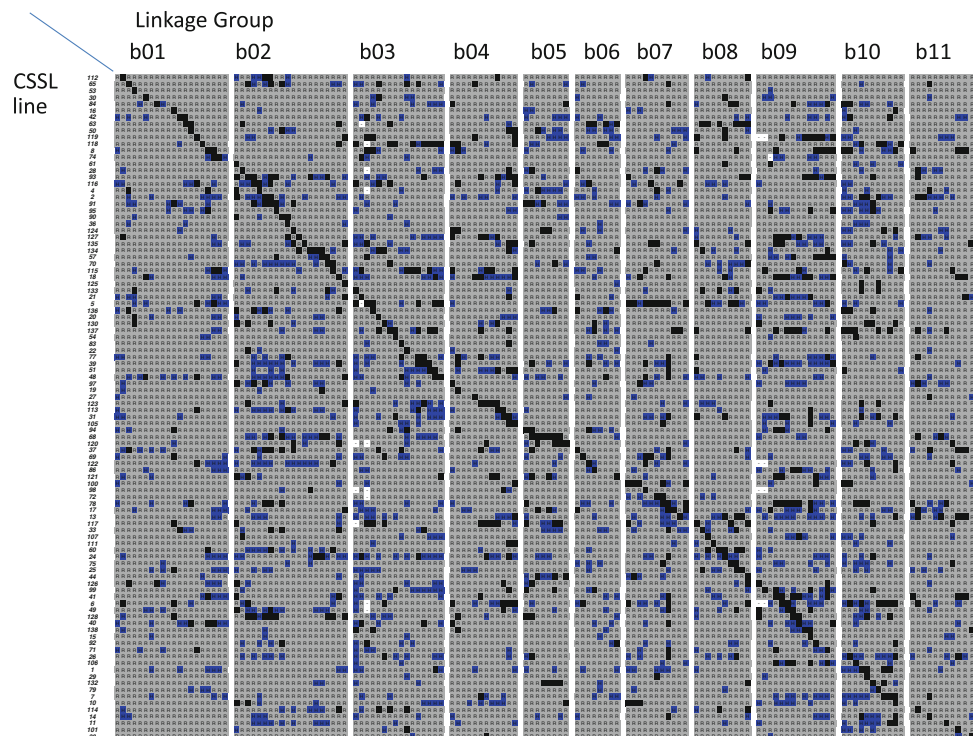


Fig. 2 Introgression line analysis for the advanced backcross population derived from the cross of (Cerinza \times (Cerinza \times (Cerinza \times G10022))). Chromosome segment substitution line (CSSL) analysis was carried out to order the introgression lines based on the most non-overlapping segments fixed for the donor parent (G10022)

allele shown in *black*. Other segment colors represent loci fixed for the recurrent parent alleles shown in *light gray* or *light blue* and heterozygous segments shown in *dark gray* or *dark blue*. Segments having missing data for a genotype are shown in *white*

mostly associated with QTL also found in Palmira. The positive alleles were always from the recurrent cultivated parent Cerinza.

In comparison, net iron content was associated with 13 markers also on nine linkage groups, with b02 markers being predominant and the only association in Darién being for the b08 marker. The positive alleles were always contributed by the higher seed weight parent, Cerinza. Zinc and iron seed content were highly correlated ($r = 0.93$ and $r = 0.88$ in Darién and Palmira, respectively) due to the affect of the distribution of seed size in the advanced backcross population and the autocorrelation of seed size in both environments and this could have affected the detection of overlapping QTL for the two traits.

The large number of SPA-QTL for mineral content is likely due to the association and correlation with seed size which also had a large number of positively associated markers. A total of 21 markers were associated with seed size in Palmira and 15 with seed size in Darién. More variability for seed size in Palmira explains the larger number of QTL detected there than in Darién. In general, Darién (1,500 masl) is considered to be a more favorable site for Andean beans like Cerinza than Palmira (1,000 masl) due to its cooler, higher elevation environment.

The QTL for seed size found with SPA analysis were located on every linkage group except for b10 and b11 with alignment between the marker associations across the two sites, especially for BM156, PV78, PV243 and BMc347 on b02, for BM98, ATA26, BM197 and PV109 on b03, BM155 on b05, PV172 and BMc316 on b08 and PV60 on b09. Interestingly, linkage group b07 was of limited importance for seed size QTL in the wild \times cultivated cross, although the previous studies have shown a role for the *Phs* locus in controlling seed size in Andean \times Mesoamerican cultivated crosses (Blair et al. 2009a).

In the composite interval mapping analysis, meanwhile, a total of 16 QTL were identified for the traits registered in the two locations (Table 3). Of these, the majority (9) were for seed weight while only one QTL each was found for yield and iron concentration in Palmira. The iron concentration QTL *Fe7.1* was located on b07 near the marker PV167 and had a LR value of 18.2. The yield QTL was found on linkage group b05 and had a LR value of 18.1. Four zinc content QTL were found (three in Palmira and one in Darién), with LR values ranging from 15.3 to 29.2, but no QTL were found for zinc concentration by CIM analysis in either location. Finally, one iron content QTL was found for grain produced in Darién. This QTL was located on b08 near the marker BMc316 and had an LR

Table 2 Quantitative trait loci for seed yield, seed mineral concentration or content and seed weight detected with single marker regression analysis for the (Cerinza × (Cerinza × (Cerinza × G10022))) advanced backcross population

| Location | LG | Marker | Significance | | | Source |
|----------------------|----|----------|--------------|----------------|--------|---------|
| | | | F | R ² | P | |
| Yield (kg/ha) | | | | | | |
| Palmira | 1 | PV233 | 6.32 | 0.10 | 0.0025 | Cerinza |
| Palmira | 1 | PV54 | 9.33 | 0.13 | 0.0002 | Cerinza |
| Palmira | 1 | PV107 | 8.5 | 0.12 | 0.0003 | Cerinza |
| Palmira | 1 | PV-atcc3 | 13.31 | 0.17 | 0.0000 | Cerinza |
| Palmira | 2 | PV78 | 8.14 | 0.11 | 0.0005 | Cerinza |
| Palmira | 3 | PV109 | 16.36 | 0.20 | 0.0000 | Cerinza |
| Palmira | 4 | BMc284 | 6.71 | 0.10 | 0.0017 | Cerinza |
| Palmira | 9 | BM202 | 7.35 | 0.10 | 0.0009 | Cerinza |
| Palmira | 9 | PV60 | 7.68 | 0.11 | 0.0007 | Cerinza |
| Palmira | 9 | PV149 | 7.51 | 0.11 | 0.0008 | Cerinza |
| Seed weight (g/100s) | | | | | | |
| Darién | 2 | BM156 | 7.35 | 0.10 | 0.0009 | Cerinza |
| Darién | 2 | PV78 | 8.69 | 0.12 | 0.0003 | Cerinza |
| Darién | 2 | PV243 | 6.79 | 0.10 | 0.0016 | Cerinza |
| Darién | 2 | BMc347 | 9.25 | 0.13 | 0.0002 | Cerinza |
| Darién | 3 | ATA26 | 6.45 | 0.10 | 0.0022 | Cerinza |
| Darién | 3 | BM197 | 7.24 | 0.10 | 0.0010 | Cerinza |
| Darién | 3 | PV109 | 13.94 | 0.18 | 0.0000 | Cerinza |
| Darién | 4 | PV-ag4 | 8.06 | 0.11 | 0.0005 | Cerinza |
| Darién | 5 | BM155 | 11.43 | 0.15 | 0.0000 | Cerinza |
| Darién | 6 | BMc371 | 9.98 | 0.15 | 0.0001 | Cerinza |
| Darién | 7 | BMd40 | 8.19 | 0.11 | 0.0004 | Cerinza |
| Darién | 8 | PV173 | 8.66 | 0.12 | 0.0003 | Cerinza |
| Darién | 8 | BMc316 | 6.49 | 0.19 | 0.0021 | Cerinza |
| Darién | 9 | PV60 | 7.66 | 0.11 | 0.0007 | Cerinza |
| Darién | 9 | ME1 | 9.03 | 0.12 | 0.0002 | Cerinza |
| Darién | 9 | BMd21 | 8.01 | 0.11 | 0.0005 | Cerinza |
| Palmira | 1 | PV-atcc3 | 11.67 | 0.15 | 0.0000 | Cerinza |
| Palmira | 2 | BM156 | 10.38 | 0.13 | 0.0001 | Cerinza |
| Palmira | 2 | PV78 | 12.96 | 0.17 | 0.0000 | Cerinza |
| Palmira | 2 | PV243 | 8.19 | 0.12 | 0.0005 | Cerinza |
| Palmira | 2 | BMd19 | 7 | 0.10 | 0.0013 | Cerinza |
| Palmira | 2 | PV-cct1 | 7.65 | 0.15 | 0.0009 | Cerinza |
| Palmira | 2 | BMc347 | 8.88 | 0.13 | 0.0003 | Cerinza |
| Palmira | 3 | BM98 | 7.65 | 0.10 | 0.0007 | Cerinza |
| Palmira | 3 | ATA26 | 6.43 | 0.10 | 0.0023 | Cerinza |
| Palmira | 3 | BM197 | 8.63 | 0.12 | 0.0003 | Cerinza |
| Palmira | 3 | PV109 | 16.62 | 0.20 | 0.0000 | Cerinza |
| Palmira | 4 | BMc284 | 6.79 | 0.10 | 0.0016 | Cerinza |
| Palmira | 5 | BM155 | 7.78 | 0.10 | 0.0006 | Cerinza |
| Palmira | 6 | BMc371 | 7.13 | 0.11 | 0.0012 | Cerinza |
| Palmira | 8 | PV173 | 7.11 | 0.10 | 0.0012 | Cerinza |
| Palmira | 8 | BMc316 | 8.33 | 0.13 | 0.0004 | Cerinza |
| Palmira | 8 | PV53 | 7.06 | 0.10 | 0.0012 | Cerinza |
| Palmira | 9 | PV60 | 7.26 | 0.11 | 0.0011 | Cerinza |

Table 2 continued

| Location | LG | Marker | Significance | | | Source |
|--------------------------|----|----------|--------------|------|--------|---------|
| | | | F | R2 | P | |
| Palmira | 9 | ME1 | 7.11 | 0.10 | 0.0012 | Cerinza |
| Palmira | 9 | BMd21 | 7.5 | 0.10 | 0.0008 | Cerinza |
| Palmira | 9 | PV149 | 8.55 | 0.13 | 0.0003 | Cerinza |
| Fe concentration (ppm) | | | | | | |
| Palmira | 10 | BMd42 | 5.71 | 0.08 | 0.0042 | Cerinza |
| Darién | 1 | PV233 | 6.51 | 0.10 | 0.0021 | G10022 |
| Darién | 1 | PV-atcc1 | 6.22 | 0.08 | 0.0026 | G10022 |
| Darién | 11 | ATA145 | 5.76 | 0.08 | 0.0041 | G10022 |
| Darién | 11 | BMc368 | 9.31 | 0.12 | 0.0002 | G10022 |
| Zinc concentration (ppm) | | | | | | |
| Palmira (ICP) | 10 | GAT11b | 5.44 | 0.08 | 0.0054 | Cerinza |
| Palmira | 10 | ATA71 | 5.36 | 0.08 | 0.0058 | G10022 |
| Palmira | 10 | BMc5 | 8.4 | 0.11 | 0.0004 | Cerinza |
| Palmira | 11 | ATA145 | 9.28 | 0.13 | 0.0002 | G10022 |
| Darién | 1 | PV31 | 6.11 | 0.08 | 0.0029 | Cerinza |
| Darién | 4 | PV-at3 | 7.11 | 0.10 | 0.0012 | G10022 |
| Darién | 11 | ATA145 | 11.62 | 0.16 | 0.0000 | G10022 |
| Darién | 11 | BMc368 | 5.62 | 0.08 | 0.0045 | G10022 |
| Iron content (mg/seed) | | | | | | |
| Palmira | 1 | PV-atcc3 | 11.87 | 0.15 | 0.0000 | Cerinza |
| Palmira | 2 | BM156 | 7.74 | 0.10 | 0.0007 | Cerinza |
| Palmira | 2 | PV78 | 9.45 | 0.13 | 0.0002 | Cerinza |
| Palmira | 2 | PV243 | 6.54 | 0.10 | 0.0020 | Cerinza |
| Palmira | 2 | BMc347 | 7.83 | 0.11 | 0.0006 | Cerinza |
| Palmira | 3 | BM197 | 7.49 | 0.10 | 0.0008 | Cerinza |
| Palmira | 3 | PV109 | 11.83 | 0.15 | 0.0000 | Cerinza |
| Palmira | 5 | BM155 | 7.09 | 0.10 | 0.0012 | Cerinza |
| Palmira | 6 | BMc371 | 7.53 | 0.11 | 0.0008 | Cerinza |
| Palmira | 7 | BMd40 | 7.97 | 0.11 | 0.0005 | Cerinza |
| Palmira | 9 | BMd21 | 8.05 | 0.11 | 0.0005 | Cerinza |
| Palmira | 10 | BMc5 | 7.88 | 0.11 | 0.0006 | Cerinza |
| Darién | 8 | ATA247 | 9.21 | 0.13 | 0.0002 | Cerinza |
| Zinc content (mg/seed) | | | | | | |
| Palmira | 1 | PV-atcc3 | 10.31 | 0.14 | 0.0001 | Cerinza |
| Palmira | 2 | BM156 | 8 | 0.11 | 0.0005 | Cerinza |
| Palmira | 2 | PV78 | 9.35 | 0.13 | 0.0002 | Cerinza |
| Palmira | 2 | PV-cct1 | 5.37 | 0.11 | 0.0063 | Cerinza |
| Palmira | 2 | BMc347 | 6.8 | 0.10 | 0.0016 | Cerinza |
| Palmira | 3 | BM98 | 7.66 | 0.10 | 0.0007 | Cerinza |
| Palmira | 3 | ATA26 | 7.21 | 0.11 | 0.0011 | Cerinza |
| Palmira | 3 | BM197 | 9.66 | 0.13 | 0.0001 | Cerinza |
| Palmira | 3 | PV109 | 16.33 | 0.20 | 0.0000 | Cerinza |
| Palmira | 4 | BMc284 | 8.13 | 0.12 | 0.0005 | Cerinza |
| Palmira | 4 | PV-ag4 | 9.03 | 0.12 | 0.0002 | Cerinza |
| Palmira | 5 | BM155 | 8.42 | 0.11 | 0.0004 | Cerinza |
| Palmira | 6 | BMc371 | 9.72 | 0.14 | 0.0001 | Cerinza |
| Palmira | 7 | BMd40 | 9.78 | 0.13 | 0.0001 | Cerinza |

Table 2 continued

| Location | LG | Marker | Significance | | | Source |
|----------|----|--------|--------------|------|--------|---------|
| | | | F | R2 | P | |
| Palmira | 8 | BMc316 | 6.4 | 0.10 | 0.0023 | Cerinza |
| Palmira | 9 | PV60 | 7.6 | 0.11 | 0.0008 | Cerinza |
| Palmira | 9 | BMd21 | 7.49 | 0.10 | 0.0008 | Cerinza |
| Palmira | 10 | BMc5 | 8.2 | 0.11 | 0.0004 | Cerinza |
| Darién | 9 | PV149 | 7.8 | 0.11 | 0.0007 | Cerinza |

Table 3 Quantitative trait loci for seed yield, seed mineral concentration and seed weight identified by composite interval (CIM) mapping analysis for the Cerinza × (Cerinza × (Cerinza × G10022)) advanced backcross population

| QTL | Location | LG | Significance value | | | Source | Additivity | Nearest marker ^a | Flanking markers ^b |
|---------------------|----------|----|--------------------|------|------|---------|------------|-----------------------------|-------------------------------|
| | | | LR | R2 | TR2 | | | | |
| Yield | | | | | | | | | |
| <i>Yld5.1</i> | Palmira | 5 | 18.13 | 0.08 | 0.51 | Cerinza | 137.6 | BM175 | BMd28–PV93 |
| Seed weight | | | | | | | | | |
| <i>Sw2.1</i> | Darién | 2 | 27.23 | 0.14 | 0.67 | Cerinza | 6.477 | BMc347 | BM152–PV-cct1 |
| <i>Sw3.1</i> | Darién | 3 | 40.39 | 0.16 | 0.74 | Cerinza | 5.700 | PV109 | PV87 |
| <i>Sw5.1</i> | Darién | 5 | 39.26 | 0.13 | 0.64 | Cerinza | 6.973 | BM155 | BMd20–BMd28 |
| <i>Sw9.1</i> | Darién | 9 | 27.98 | 0.11 | 0.67 | Cerinza | 4.419 | BMd21 | BMd54–BM148 |
| <i>Sw9.2</i> | Darién | 9 | 21.59 | 0.08 | 0.65 | Cerinza | 4.168 | ATA9 | BM148–BM154 |
| <i>Sw2.2</i> | Palmira | 2 | 25.62 | 0.12 | 0.66 | Cerinza | 6.194 | PV78 | PV243–BM156 |
| <i>Sw3.2</i> | Palmira | 3 | 27.92 | 0.09 | 0.61 | Cerinza | 4.906 | PV109 | PV87 |
| <i>Sw5.2</i> | Palmira | 5 | 33.12 | 0.12 | 0.62 | Cerinza | 6.970 | BM155 | BMd20–BMd28 |
| <i>Sw5.3</i> | Palmira | 5 | 26.08 | 0.11 | 0.61 | Cerinza | 6.299 | BMd28 | BM155–BM175 |
| Iron conc. | | | | | | | | | |
| <i>Fe7.1</i> | Palmira | 7 | 18.24 | 0.08 | 0.40 | G10022 | −4.762 | PV167 | PV35–BMc339 |
| Iron content | | | | | | | | | |
| <i>Fe_cont8.1</i> | Darién | 8 | 16.86 | 0.12 | 0.34 | Cerinza | 0.408 | BMc316 | BMc258–PV173 |
| Zinc content | | | | | | | | | |
| <i>Zn_cont2.1</i> | Palmira | 3 | 29.24 | 0.12 | 0.58 | Cerinza | 0.185 | PV109 | PV87 |
| <i>Zn_cont5.1</i> | Palmira | 5 | 32.69 | 0.13 | 0.57 | Cerinza | 0.241 | BM155 | BMd20–BMd28 |
| <i>Zn_cont5.2</i> | Palmira | 5 | 25.84 | 0.12 | 0.56 | Cerinza | 0.216 | BMd28 | BM155–BM175 |
| <i>Zn_cont7.1</i> | Darién | 7 | 15.34 | 0.10 | 0.44 | G10022 | −0.156 | PV35 | PV242–PV167 |

Values have significance at 0.5 % probability after 1000-fold permutation tests

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

^a The 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

value of 16.8. The seed weight QTL was detected in both Darién (5 QTL) and Palmira (4 QTL) across four linkage groups (b02, b03, b05 and b09) and were generally of high LR values (ranging up to 40.4) compared to the QTL for other traits.

Among these QTL, there was overlap for the QTL identified in Darién and Palmira on linkage groups b03 and b05 showing that G×E effects for this trait were low. The exact locations of the QTL and their overlaps are shown in

Fig. 3. All the QTL detected by CIM analysis for seed weight and iron content showed positive effects with the alleles from the recurrent parent, Cerinza. Meanwhile the alleles from G10022 showed positive affects for the yield QTL *Yld5.1*, the iron concentration QTL *Fe7.1* and for the zinc content QTL *Zn_cont7.1*. The last of these two QTL were located in the same region and would be present on the same wild donor parent introgressions for the most part. In general, the QTL for seed mineral

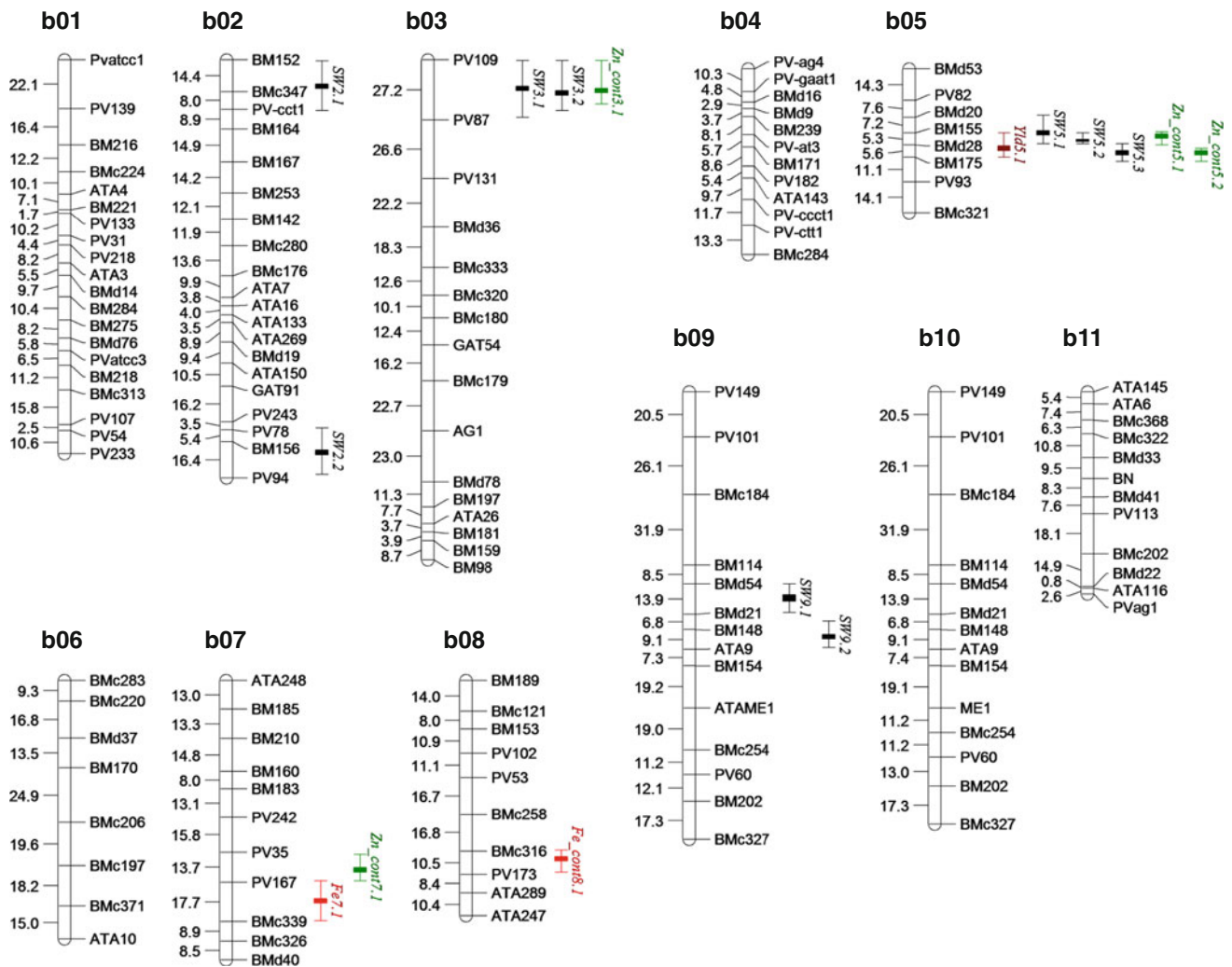


Fig. 3 Genetic linkage map for the advanced backcross inbred line population developed from the cross the (Cerinza × (Cerinza × (Cerinza × G10022))) advanced backcross population showing location of quantitative trait loci for yield, seed weight and

mineral concentration and content traits. QTL are identified as vertical bars crossed with a horizontal bar at the peak LR value. *Fe* iron, *Zn* zinc, *Sw* 100 seed weight and *Yld* yield

content in the CIM analysis did not align with those of the SPA analysis. However, for the seed weight QTL some overlap was observed between the QTL detection techniques.

Discussion

In our analysis of the wild × cultivated advanced backcross population generated from a commercial Andean variety Cerinza with a very high seed mineral wild genotype G10022, the inheritance of iron and zinc accumulation was quantitative and many QTL controlled iron and zinc concentrations and contents per seed. Similar results were observed in the previous study of seed mineral accumulation in a wild × cultivated recombinant inbred line cross

(Guzmán-Maldonado et al. 2003) and in all the previous studies of cultivated common bean populations (Blair et al. 2009a, 2010b, 2011a; Cichy et al. 2009). It appears that in certain commercial classes, such as navy beans that are sensitive to zinc-deficient soils, inheritance may be more simple (Gelin et al. 2007) but in general all the studies of inheritance have shown oligogenic inheritance of seed mineral traits with small genotype × year effects and higher genotype × location effects.

Alignment of the QTL locations from the present study and from the previous studies by Blair et al. (2009a, 2010b, 2011a) and Cichy et al. (2009) was made possible by the anchored and syntenic genetic map used in each of these studies. Central to map anchoring and QTL alignment was the use of locus-specific SSR markers that were common to several of the populations as well as full genome coverage

to avoid missing comparable QTL locations through limited map saturation.

In the first comparison, we expected the QTL identified in this study to be similar to those identified for the Andean \times Mesoamerican population in Blair et al. (2009a). Comparisons showed QTL alignment especially for CIM iron concentration QTL detected in this study on linkage groups b07 and b08 with similarly positioned CIM-detected QTL on the DOR364 \times G19833 map. Interestingly, the QTL were found with both ICP and AAS analytical methods in the study from Blair et al. (2009a) and the R² values were similar (9 %). Meanwhile, no QTL for zinc concentration with CIM analysis were found in this study but zinc content QTL on linkage groups b03 and b07 were found near zinc concentration QTL from the DOR364 \times G19833 study on linkage groups b03 and b07. Single point analysis QTL found on linkage group b01, b06 and b11 for iron concentration and on b11 for zinc concentration in the present study could confirm QTL for these minerals found in Blair et al. (2009a).

Comparisons of the QTL found in this study and in the DOR364 \times G19833 population with those found in Andean \times Andean populations from Blair et al. (2011a) and Cichy et al. (2009) are interesting because of the QTL that might be genepool specific and not detectable in inter-genepool crosses. We found specific QTL on linkage groups b02 and b09 for iron concentration and b02 and b08 for zinc concentration. Cichy et al. (2009) found that in determinate \times indeterminate bush bean Andean crosses, the region of the linkage group b01 near the *fin* locus had an effect on iron and zinc concentration which were not found. However, some overlap did occur with iron QTL from both Blair et al. (2009a) and Cichy et al. (2009) on linkage groups b05, b06, b08, b09 and b11 where overlapping zinc concentration QTL were also found. A similar comparison for the Mesoamerican \times Mesoamerican population analyzed by Blair et al. (2010b) versus the inter-genepool crosses of Cerinza \times G10022 and DOR364 \times G19833, showed that b04 and b06 linkage group QTL for iron concentration, b01 and b07 for iron content and b01 for zinc content may be Mesoamerican genepool specific. Meanwhile, linkage group b03 QTL for zinc concentration may align with zinc content QTL from Blair et al. (2010b).

The smaller number of QTL identified, in the present CIM analysis compared to the cultivated \times cultivated inter-genepool population analyzed by Blair et al. (2009a) or the intra-genepool populations of Blair et al. (2010b, 2011a) and Cichy et al. (2009), may be due to the intermediate seed mineral accumulation of Cerinza. It was notable, that Cerinza was among the higher seed mineral cultivars among 40 varieties analyzed by Astudillo and Blair (2008). In that analysis, Cerinza had 53.7 and 60.6 ppm iron concentration and 27.6 and 24.0 ppm zinc

concentration under low and high soil phosphorus levels, respectively. In Blair et al. (2005), Cerinza had even higher seed mineral values of 78 ppm iron and 35 ppm zinc. In the study here the comparison of seed mineral accumulation in Palmira and Darién shows the G \times E effects to be important for soil types in these two sites. Therefore, some of the differences between the parents in this study can be attributed to cultivated versus wild status while others could be due to G \times E effects and genepool differences. Notably, Cerinza is from the Andean genepool, which in general has higher iron but lower zinc than genotypes from the Mesoamerican genepool (Blair et al. 2009a). Meanwhile, seed of the wild parent G10022 is a wild Mesoamerican bean known to have 108 ppm seed iron and 38 ppm seed zinc.

Another explanation for lower than expected number of seed mineral QTL from the wild parent alleles detectable in the advanced backcross population is that segregation distortion was significant in the population. Higher than expected frequency of the recurrent parent alleles was observed across the genome. To a certain extent, the domestication genes described by Koinange et al. (1996), could explain the patterns of segregation distortion as was observed in another common bean advanced backcross population analyzed by Blair et al. (2006b). In that case, the pattern of segregation distortion in the advanced backcross population created here for G10022 was similar to that of the advanced backcross population created for G24404. In both cases, segregation distortion was found for a large number of markers on b01, b02, b08 and b09.

In the earlier population from Blair et al. (2006b), distortion was also mostly against the wild alleles and toward a predominance of the cultivated or Cerinza alleles. Since a low level of selection was made for return to recurrent parent plant architecture and seed type the predominance of the Cerinza allele at a large number of loci was to be expected. In that case, segregation distortion was high around the known map positions of genes for the domestication syndrome in common beans (Koinange et al. 1996) as well as around the architectural gene *fin* on linkage group b01 Coyne (1970). The domestication syndrome genes probably influencing selection in the (Cerinza \times (Cerinza \times (Cerinza \times G10022))) population included a gene for pod shattering on linkage group b02 and suite of seed color genes on linkage group b08.

Significantly, the *fin* locus controls determinacy in many crosses (Coyne 1967) and the selection of bush bean architecture caused skewing toward the allele from Cerinza in the population analyzed here as well as in the previous advanced backcross population. The climbing bean architecture typical of the wild parent was naturally selected against given that the population was developed without staking. The *Ppd* locus, which is involved in photoperiod

response, is also found on linkage group b01 (Gu et al. 1998) and could have been selected against as well, since the population had uniform and early flowering dates like the recurrent parent compared to the wild donor parent which is late flowering and photoperiod sensitive (data not shown).

The allele for pod shattering at the *St* locus on linkage group b02 (Koinange et al. 1996) would have been selected against as well in the advanced backcross population as shattered seed are lost in single seed descent. Finally, the seed coat color loci on linkage group b01 (Z) and b08 [R–C] or at other locations in the genome (McClellan et al. 2002) could influence segregation distortion on various chromosomes. Segregation distortion in wide crosses such as the inter-genepool population analyzed can also be due to dwarf lethal (*dl*) incompatibility genes which are known to segregate in crosses between the genepools (Hannah et al. 2007). These genes could be responsible for some of the segregation distortion in the wild × cultivated cross analyzed here.

Seed size genes are certainly part of the domestication suite as it was a priority for early agriculturalist to have larger seed both for planting and consumption. However, unlike the major domestication genes, seed size loci are of smaller penetrance and have smaller individual effects on the trait. In this study, we found that seed weight was highly consistent across sites although more variability was found for the population individuals in the moderate adaptation site of Palmira than in the high adaptation site of Darién. Despite this, QTL loci were consistent across detection methods (CIM versus SPA analysis) and across locations (both Darién and Palmira). In conclusion, seed weight was controlled as a multigenic-oligogenic trait by many loci segregating in the wild × cultivated populations.

In comparison to previous research, we did not find the phaseolin locus to be important for seed weight in the present population although it was found important for seed weight in a wild × cultivated advanced backcross by Blair et al. (2006b) and in one of the populations analyzed for micronutrients (Blair et al. 2009a). Some QTL on linkage groups b02 and b08 were consistent between the two advanced backcross populations analyzed so far in common bean for this trait. Several other QTL for seed weight may be associated with QTL for seed weight found by Tar'an et al. (2002) and Park et al. (2000), although map alignments are difficult with these RAPD-based maps. Finally, one specific QTL on linkage group b09 may be associated with one found by Vallejos and Chase (1991) on one of the early genetic maps for common bean and was also identified in the wild × cultivated cross.

It was notable that a range of seed weight QTL were detected across both locations used in this study. This

consistency of the seed weight QTL may reflect the higher heritability of this trait even though QTL analysis suggests that the trait is controlled by multiple loci. Finally, the QTL for higher yield on linkage group b05 was associated with QTL for smaller seed weight. Similar associations between yield and seed weight were found but on different linkage groups in the analysis of an advanced backcross population created with a Colombian wild bean in Blair et al. (2006b). These negative associations could be due to pleiotropy of these QTL and compensation between these traits.

As in previous studies by Blair et al. (2009a, 2010b, 2011a), Cichy et al. (2009) and Graham et al. (1999) correlations between the iron and zinc seed mineral accumulation suggest some shared physiological uptake mechanisms between the two minerals from the soil. Furthermore, zinc and iron seed contents were highly correlated ($r = 0.93$ and $r = 0.88$ in Darién and Palmira, respectively) due to the affect of the distribution of seed size in the advanced backcross population and the autocorrelation of seed size between both environments.

In conclusion, the AB-QTL method was useful for simultaneously understanding the QTL controlling the nutrition traits evaluated in this study as well as breeding for these traits into an acceptable genetic background. Although wild beans have been used before to transfer biotic stress resistance traits such as bruchid insect resistance, this study is the first to attempt to simultaneously obtain a higher seed mineral content from wild beans and to analyze QTL associated with loci controlling this traits. The factors underlying the iron concentration and content QTL will be of interest to study further. This would include candidate genes for iron deficient soil tolerance in soybean and iron uptake genes in common bean evaluated through synteny among legumes. Of special interest is the prevalence of QTL for seed iron accumulation on linkage group b07 and b08 that may have to do with important seed proteins (phaseolin and seed ferritin, respectively) which are encoded by loci on these chromosomes. QTL at these locations were also important in the cultivated × cultivated inter-genepool cross analyzed for seed iron concentration by Blair et al. (2009a) and may also be characteristic of the inheritance of seed mineral concentration in wild × cultivated crosses. Breeding for phaseolin alleles has been suggested before to increase protein content (Sullivan and Bliss 1983).

Interestingly, the regions of linkage groups b06 and b11 do not appear to contribute important QTL in crosses of wild × cultivated beans unlike in the cultivated × cultivated crosses analyzed previously (Blair et al. 2009a, 2010b, 2011a). This is perhaps due to differences in the loci segregating in wild common beans versus cultivated landraces. Certainly, the candidate genes for zinc concentration are less obvious and none to date have been well

characterized. Finally, seed iron-related QTL found in this study were not associated with iron reductase activity QTL found in DOR364 × G19833 population but this may be due to the poor root iron reductase activity of most wild accessions compared to cultivated beans (Blair et al. 2010a).

In terms of molecular breeding, the advanced backcross method is robust for QTL detection as long as large populations are used to ensure introgression from each regions of the genome even around domestication genes. The creation of CSSL isolines is an advantage for future map-based cloning of QTL and careful association of QTL with candidate genes. The advanced backcross method also had advantages in the production of selected breeding lines useful for genetic improvement of Andean and Mesoamerican beans using marker-assisted selection of introgression events from the wild accession.

Acknowledgments We are grateful 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 and to C. Astudillo for laboratory training. We are also grateful to Octavio Mosquera and Teresa Fowles for assistance with AAS and ICP analyses, respectively.

References

- Afanador LK, Hadley SD, Kelly JD (1993) Adoption of a mini-prep DNA extraction method for RAPD marker analysis in common bean. *Bean Improv Coop* 35:10–11
- Ando T, Yamamoto T, Shimizu T, Ma X, Shomura A, Takeguchi Y et al (2008) Genetic dissection and pyramiding of quantitative traits for panicle architecture by using chromosomal segment substitution lines in rice. *Theor Appl Genet* 119:881–890
- Astudillo C, Blair MW (2008) Evaluación del contenido de hierro y zinc en semilla y su respuesta al nivel de fósforo en variedades de frijol colombianas. *Agronomía Colombiana* 26:471–476
- Basten C, Weir BS, Zeng Z-B (2001) Windows QTL Cartographer. v 1.21. Software. Department of Statistics. North Carolina State University, USA
- 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
- Blair MW, Astudillo C, Restrepo J, Bravo LC, Villada D, Beebe SE (2005) Análisis multi-local de líneas de frijol arbustivo con alto contenido de hierro en el departamento de Nariño. *Fitotecnia Colombiana* 5:20–27
- Blair MW, Giraldo MC, Buendía HF, Tovar E, Duque MC, Beebe SE (2006a) Microsatellite marker diversity in common bean (*Phaseolus vulgaris* L.). *Theor Appl Genet* 113:100–109
- Blair MW, Iriarte G, Beebe S (2006b) QTL analysis of yield traits in an advanced backcross population derived from a cultivated Andean × wild common bean (*Phaseolus vulgaris* L.) cross. *Theor Appl Genet* 112:1149–1163
- Blair MW, Buendía HF, Giraldo MC, Métails I, Peltier D (2008) Characterization of AT-rich microsatellites in common bean (*Phaseolus vulgaris* L.). *Theor Appl Genet* 118:91–103
- Blair MW, Astudillo Beebe SE, Roa I, Kimani P, Chirwa R (2009a) Biofortification of common bean (*Phaseolus vulgaris* L.) via traditional and novel breeding approaches. *J Danish Biochem Soc (Biozoom)* 1:25–28
- Blair MW, Astudillo C, Grusak M, Graham R, Beebe S (2009b) Inheritance of seed iron and zinc content in common bean (*Phaseolus vulgaris* L.). *Mol Breeding* 23:197–207
- Blair MW, Muñoz M, Pedraza F, Giraldo MC, Buendía HF, Hurtado N (2009c) Development of microsatellite markers for common bean (*Phaseolus vulgaris* L.) based on screening of non-enriched small insert genomic libraries. *Genome* 52:772–782
- Blair MW, Muñoz-Torres M, Giraldo MC, Pedraza F (2009d) Development and diversity assessment of Andean-derived, gene-based microsatellites for common bean (*Phaseolus vulgaris* L.). *BMC Plant Bio* 9:100
- Blair MW, Knewton SJB, Astudillo C, Li CM, Fernandez AC, Grusak MA (2010a) Variation and inheritance of iron reductase activity in the roots of common bean (*Phaseolus vulgaris* L.) and association with seed iron accumulation QTL. *BMC Plant Biol* 10:215
- Blair MW, Medina JJ, Astudillo C, Rengifo J, Beebe SE, Machado G, Graham R (2010b) QTL for seed iron and zinc concentrations in a recombinant inbred line population of Mesoamerican common beans (*Phaseolus vulgaris* L.). *Theor Appl Genet* 121:1059–1071
- Blair MW, Astudillo C, Rengifo J, Beebe SE, Graham R (2011a) QTL for seed iron and zinc concentrations in a recombinant inbred line population of Andean common beans (*Phaseolus vulgaris* L.). *Theor Appl Genet* 122:511–523
- Blair MW, Hurtado N, Chavarro MC, Muñoz-Torres MC, Tomkins J, Wing R (2011b) Gene-based SSR markers for common bean (*Phaseolus vulgaris* L.) derived from root and leaf tissue ESTs: an integration of the BMc series. *BMC Plant Bio* 11:50
- Broughton WJ, Hernandez G, Blair MW, Beebe SE, Gepts P, Vanderleyden J (2003) Beans (*Phaseolus* spp.)—model food legumes. *Plant Soil* 252:55–128
- Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. *Genetics* 138:963–971
- Cichy KA, Caldas GV, Snapp SS, Blair MW (2009) QTL analysis of seed iron, zinc, and phosphorus levels in an Andean bean population. *Crop Sci* 49:1742–1750
- 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
- Delaney DE, Bliss EA (1991) Selection for increased percentage phaseolin in common bean. 2. Changes in frequency of seed protein alleles with S1 family recurrent selection. *Theor Appl Genet* 93:306–311
- Eshed Y, Zamir D (1995) An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield associated QTL. *Genetics* 139:1147–1162
- Fulton T, Emmatty D, Eshed Y, Lopez J, Petiard V et al (1997) QTL analysis of an advanced backcross of *Lycopersicon peruvianum* to the cultivated tomato and comparisons with QTLs found in other wild species. *Theor Appl Genet* 114:881–894
- Gelin JR, Forster S, Grafton KF, McClean P, Rojas-Cifuentes GA (2007) Analysis of seed-zinc and other nutrients in a recombinant inbred population of navy bean (*Phaseolus vulgaris* L.). *Crop Sci* 47:1361–1366
- Ghesquiere A, Sequier J, Second G, Lorieux M (1997) First steps toward a rational use of African rice, *Oryza glaberrima*, in rice breeding: a contig line concept. *Euphytica* 101:31–39
- Graham R, Senadhira D, Beebe S, Iglesias C, Monasterio I (1999) Breeding for micronutrient density in edible portions of staple food crops: conventional approaches. *Field Crops Res* 60:57–80

- Grisi MCM, Blair MW, Gepts P, Brondani C, Pereira PAA, Brondani RPV (2007a) Genetic mapping of microsatellite markers in common bean (*Phaseolus vulgaris*) population BAT93 × Jalo EEP558. *Genet Mol Res* 6:691–706
- Grisi MCM, Blair MW, Gepts P, Brondani C, Pereira PAA, Brondani RPV (2007b) Genetic mapping of microsatellite markers in common bean (*Phaseolus vulgaris*) population BAT93 × Jalo EEP558. *Genet Mol Res* 6:691–706
- 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
- Guzmán-Maldonado SH, Martínez O, Acosta-Gallegos JA, Guevara-Lara F (2003) Putative quantitative trait loci for physical and chemical components of common bean. *Crop Sci* 43:1029–1035
- Guzmán-Maldonado SH, Acosta-Gallegos J, Paredes-López O (2004) Protein and mineral content of a novel collection of wild and weedy common bean (*Phaseolus vulgaris* L.). *J Sci Food Agric* 80:1874–1881
- Hannah MA, Krämer KM, Geffroy V, Kopka J, Blair MW, Erban A, Vallejos CE, Heyer AG, Sanders FET, Millner PA, Pilbeam DJ (2007) The DL gene system in common bean (*Phaseolus vulgaris* L.) causes programmed root death due to a shoot-derived inhibitory signal. *New Phytol* 176:537–549
- Islam FMA, Basford KE, Jara C, Redden RJ, Beebe SE (2002) Seed compositional and disease resistance differences among gene pools in cultivated common bean. *Genet Resour Crop Evol* 49:285–293
- Koinange EMK, Singh SP, Gepts P (1996) Genetic control of the domestication syndrome in common bean. *Crop Sci* 36:1037–1045
- 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
- Lin S, Sasaki T, Yano M (1998) Mapping quantitative trait loci controlling seed dormancy and heading date in rice, *Oryza sativa*, using backcross inbred lines. *Theor Appl Genet* 100:997–1003
- 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, Jalouzet 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
- Park SO, Coyne DP, Jung G, Skroch PW, Arnaud-Santana E, Steadman JR, Ariyaratne HM, Nienhuis J (2000) Mapping of QTL for seed size and shape traits in common bean. *J Amer Soc Hort Sci* 125:466–475
- Pfeiffer W, McClafferty B (2007) HarvestPlus: breeding crops for better nutrition. *Crop Sci* 47:S88–S105
- Singh SP (2001) Broadening the genetic base of common bean cultivars: a review. *Crop Sci* 41:659–1675
- 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
- Vallejos CE, Chase C (1991) Extended map for the phaseolin linkage group of *Phaseolus vulgaris* L. *Theor Appl Genet* 87:353–357
- Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered* 93:77–78
- Wortmann CS, Kirkby RA, Eledu CA, Allen DJ (1998) Atlas of common bean (*Phaseolus vulgaris* L.) production in Africa. Pan-African Bean Research Alliance, p 133
- 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