Theor Appl Genet (2004) 108:243–252
DOI 10.1007/s00122-003-1437-3

ORIGINAL PAPER

F. M. Amirul Islam · S. Beebe · M. Muñoz · J. Tohme · R. J. Redden · K. E. Basford

Using molecular markers to assess the effect of introgression on quantitative attributes of common bean in the Andean gene pool

Received: 28 January 2003 / Accepted: 13 August 2003 / Published online: 5 December 2003

Abstract Progress in bean breeding programs requires the exploitation of genetic variation that is present among races or through introgression across gene pools of *Phaseolus vulgaris* L. Of the two major common bean gene pools, the Andean gene pool seems to have a narrow genetic base, with about 10% of the accessions in the CIAT core collection presenting evidence of introgression. The objective of this study was to quantify the degree of spontaneous introgression in a sample of common bean landraces from the Andean gene pool. The effects of introgression on morphological, economic and nutritional attributes were also investigated. Homogeneity analysis was performed on molecular marker data from 426 Andean-type accessions from the primary centres of origin of the CIAT common bean core collection and two check varieties. Quantitative attribute diversity for 15 traits was studied based on the groups found from the cluster analysis of marker prevalence indices computed for each accession. The two-group summary consisted of one group of 58 accessions (14%) with low prevalence indices and another group of 370 accessions (86%) with high prevalence indices. The smaller group occupied the outlying area of points displayed from homogeneity analysis, yet their geographic origin was widely distributed over the Andean region. This group was regarded as introgressed, since its accessions displayed traits that are associated with the Middle American gene pool: high resistance to Andean disease isolates but low resistance to Middle American disease isolates, low seed weight and high scores for all nutrient elements. Genotypes generated by spontaneous introgression can be helpful for breeders to overcome the difficulties in transferring traits between gene pools.

Communicated by H.C. Becker

F. M. Amirul Islam
Western Australian Institute for Medical Research,
Australia, 6009 Nedlands, WA, Australia

S. Beebe · M. Muñoz · J. Tohme
Centro Internacional de Agricultura Tropical (CIAT),
Cali, Colombia

R. J. Redden
Department of Natural Resources and Environment,
Agriculture Victoria, 3401 Horsham, Vic, Australia

K. E. Basford (✉)
School of Land and Food Sciences,
The University of Queensland, 4072 Brisbane, Qld, Australia
e-mail: k.e.basford@uq.edu.au

Introduction

The common bean (*Phaseolus vulgaris* L.) consists of two major gene pools, one Middle American and one Andean (Gepts 1988a), as determined by morphological attributes (Singh et al. 1991a, 1991b), seed proteins (Gepts 1988b), isozymes (Koenig and Gepts 1989) and DNA markers (Khairallah et al. 1990; Gepts and Debouck 1991; Becerra Velásquez and Gepts 1994). The Middle American gene pool is characterized by accessions with small and medium seed, mostly S, Sb, Sd and M phaseolin, while the Andean gene pool is typified by medium to large seed and mostly T, H and C phaseolin (Singh et al. 1991a, 1991b; Beebe et al. 2000). These gene pools reflect multiple domestication events within distinct wild populations in the respective regions (Gepts and Debouck 1991). In addition to these two major gene pools, Islam et al. (2002a, 2002b) identified a minor North Andean gene pool.

Gene pools appear to have co-evolved with local pathogen populations, such that accessions in one gene pool are often susceptible to pathogen races from the same region but relatively more resistant to those from the other region. This phenomenon has been observed in the case of *Colletotrichum lindemuthianum*, the anthracnose pathogen (Balardin and Kelly 1998; Geffroy et al. 1999), and *Phaeoisariopsis griseola*, the angular leaf spot pathogen (Guzman et al. 1995; Pastor-Corrales et al. 1998; Islam et al. 2002c).

The Middle American gene pool is composed of distinct races (Singh et al. 1991a) that can be distin-

guished at the DNA level using random amplified polymorphic DNA (RAPD) markers (Beebe et al. 2000). Race Durango consists of both prostrate and climbing beans and originates in the dry plateau and highlands of Mexico, while races Jalisco and Guatemala are composed of climbing beans that exist in the moist mountainous regions. The small-seeded Middle American race is typical of the lowland tropics and, in turn, can be subdivided (on the basis of phenotype and RAPD analysis) into sub-race M1, which is largely of upright type-II indeterminate growth habit and is localized in tropical Mexico, and sub-race M2, which is of indeterminate prostrate type-III habit and is more common in Central America (Beebe et al. 2000). Accessions of sub-race M2 have also been found in the Andean Zone (Beebe et al. 2000), and the archaeological record suggests that these have been in the Andean region since pre-Colombian times (Towle 1961; Kaplan and Kaplan 1988).

The Andean gene pool has been divided into three races based on growth habit and adaptation regimes (Singh et al. 1991a). Race Nueva Granada is represented by determinate and indeterminate bush types, and race Peru is composed of climbing beans adapted to high altitudes. Race Chile includes indeterminate prostrate types adapted to high latitudes. However, DNA analysis indicated that the Andean gene pool has a very narrow genetic base and that the races were not distinguishable based on DNA analysis with amplified fragment length polymorphic (AFLP) markers (Beebe et al. 2001). DNA analysis of wild bean populations suggested that the eastern slopes of the Andes, especially in Bolivia or Argentina, could have been a primary domestication site of Andean common bean (Beebe et al. 2001). Kami et al. (1995) studied the DNA sequence of the phaseolin gene from Ecuadorean wild beans and proposed that this population could be ancestral to both major gene pools.

Progress in breeding requires the exploitation of genetic variation among races and gene pools. Singh and Gutiérrez (1990) reported yield gains in interracial crosses within the Middle American gene pool, however there has been slower progress in the improvement of Andean types (Kornegay et al. 1992; White et al. 1992). The narrow genetic base among Andean types may be a limitation on genetic gain when intra-gene pool crosses are made (Beebe et al. 2001). One strategy for the improvement of Andean beans is through crossing with Middle American beans, although these crosses have often not been productive (Singh 1995). Some extra effort, such as recurrent selection, is necessary to obtain useful progeny from intergene pool crosses (Beaver and Kelly 1994; Singh et al. 1999). Barriers to inter-gene pool crosses, including F_1 hybrid lethality, have been observed (Gepts and Bliss 1985), and while these extreme cases are not common, they may be indicative of other underlying genetic differences that obstruct the exploitation of variability across gene pools. Surprisingly, almost 10% of the Andean accessions from the International Centre for Tropical Agriculture (CIAT) core collection exhibit evidence of introgression from Middle American beans (Beebe et al. 2001). Paredes and Gepts (1995) found an extensive introgression of Middle American germplasm into Chilean common bean cultivars. This introgression was assumed to result from spontaneous outcrossing in farmers' fields, based on segregation found previously in farmers' varietal mixtures (Beebe et al. 1997).

While introgression from the Middle American gene pool to the Andean gene pool appears to be relatively common in the Andean zones (Debouck et al. 1989; Beebe et al. 1997), its effects have not been established. The objectives of the investigation presented here were to: (1) confirm the level of introgression in the Andean gene pool and (2) quantify the effects of introgression on the phenotype of Andean landraces.

Materials and methods

Accessions

CIAT has a common bean core collection of 1,441 accessions representing both primary and secondary centres of origin. These have been selected on the basis of agroecological origins (variability in rainfall, soil, temperature and daylength at flowering time at site of origin) as well as plant and seed morphology (Tohme et al. 1995). To investigate the Andean gene, we selected 426 accessions for the present study. One criterion of selection of the accessions under study was geographic origin; thus 96% originated in the Andean countries. However, as there has been interchange of germplasm among regions for centuries, some Andean types are in fact found in the Middle American region, and vice-versa (Gepts 1988b). Since phaseolin is an important criterion of gene pool designation (Gepts 1988b), it was used both to confirm accessions from Andean countries as being potentially of the Andean gene pool, and to identify accessions from Middle America that could be of Andean origin. The geographic origin of 408 accessions was Andean (from Peru, Colombia, Ecuador, Bolivia, Chile and Argentina), while the origin of the other 18 accessions was Mexico, Guatemala and Costa Rico. In addition to these 426 accessions, two check cultivars were included for comparison: race Mesoamerica cv. ICA-Pijao, representing the Middle American gene pool, with small black seed and an upright bush indeterminate growth habit, and Andean cv. Calima of race Nueva Granada, with large red and white mottled seed and a bush determinate growth habit.

RAPD-based data

DNA was extracted from leaves of ten seedlings (bulked) of each accession and RAPD reactions carried out as per Skroch et al. (1998). Ten primers that had been selected from previous studies (Skroch et al. 1998; Beebe et al. 2000) as being especially polymorphic both within and between gene pools were run in CIAT's laboratory: AA19, AK06, H19, I07, K12, L04, O20, V10, W06 and X11, all from Operon Technologies (Alameda, Calif.). Each of 151 distinct bands observed to be polymorphic between any two accessions was scored as present or absent in each accession. A data matrix was assembled with 428 rows (accessions) and 151 columns (polymorphic bands).

Quantitative attributes

Disease reactions

Anthracnose (ANT), caused by the fungus *Colletotrichum lindemuthianum*, and angular leaf spot (ALS), caused by the fungus

Phaeoisariopsis griseola, were evaluated using Andean and Middle American pathogen isolates in independent inoculations (denoted by suffix a for Andean and m for Middle American). Isolates of the ANT pathogen included Andean races 7 and 15 and Mesoamerican races 137, 521, 385 and 1545 (Pastor-Corrales 1991), while isolates of the ALS included Andean race 63-0 and Mesoamerican races 5-47 and 31-47 (Pastor-Corrales et al.). The common bacterial blight (CBB) incited by *Xanthomonas axonopodis* pv. *phaseoli* (Smith) was evaluated on leaves. The bush-type accessions (growth habits I, II and III; Schoonhoven and Pastor-Corrales 1987) were evaluated for one insect pest, leafhopper (*Empoasca kraemeri*). The disease and pest data were described in detail by Islam et al. (2002c).

Morphological attributes

The morphological attributes were days to flower (DF), as registered when 50% of plants presented at least one open flower, days to maturity (DM), when 50% of the pods had changed color, and 100 seed weight (SWt).

Nutritional attributes

Nutritional attributes included the total percentage of seed protein in grain (PC), measured using near infra red spectrophotometry (NIRS), and mineral concentrations in the grain [calcium (Ca), phosphorus (P) and sulfur (S) as a percentage; iron (Fe) and zinc (Zn) in mg/kg]. Mineral concentration data were collected by inductively coupled plasma (ICP), a sensitive and quick method that can be used to assay a wide range of minerals (Islam et al. 2002b).

Marker prevalence in the germplasm pool

To examine the relationship between the prevalence or rarity of markers and the agronomic characteristics of the cultivars bearing those markers, we computed the frequency of each of 151 markers within the germplasm pool as the number of cultivars in which the specific marker was present divided by the number of cultivars that were scored for that marker. A 'marker prevalence index' was subsequently computed for each cultivar as the average of the frequency values of the markers present in that cultivar (Beer et al. 1997).

To illustrate this calculation, consider a simple case of four accessions with five markers. Accession A has markers 1, 2 and 3; accession B has markers 2, 3 and 5; accession C has markers 1, 2 and 5; accession D has markers 1, 2, 3 and 4. The marker ratio scores (the ratio of the markers present to the total number of accessions present in the germplasm pool) for markers 1–5 will be 0.75 (marker 1 is present three times in four accessions), 1.00, 0.75, 0.25 and 0.5, respectively. The marker prevalence index for accession A will be (0.75+1.00+0.75)/3=0.83; similarly, indices for accessions B, C and D will be 0.75, 0.75 and 0.69, respectively. The accessions with large marker prevalence indices contain markers that are widespread in the germplasm pool, while accessions with small indices contain relatively rare markers.

Analytical methods

The mixture method of clustering (Basford and McLachlan 1985) was applied to the 428×1 accession by marker prevalence index array. This methodology assumes that the accessions are selected from a mixture of normal distributions, with individual accessions able to have a non-zero probability of belonging to more than one group. It was implemented using the program EMMIX (Peel and McLachlan 1999), assuming arbitrary variances in the groups. Group membership was specified by different symbols for plotting purposes to superimpose the grouping solution on the output of the homogeneity analysis described next. The origin of the accessions within each cluster was mapped for the passport data using the software package FLORA MAP, produced by Jones et al. (1997) and Jones and Gladcov (1999).

Homogeneity analysis (Manly 1994) was used to investigate the proximity among the accessions as determined by the RAPD bands, i.e. using the raw data (0 for absent and 1 for present for each of the 151 markers). The program HOMALS, which is a part of the Category package contained in SPSS (SPSS for Windows 1999), was used for this purpose.

Quantitative trait diversity was studied based on the groups of accessions found via the clustering process on the marker prevalence index array. A simple *t*-test was performed to test the hypothesis that the means of the groups were equal for each of the quantitative traits as per Snedecor and Cochran (1989).

Results

Marker prevalence indices and the application of the mixture model

The marker prevalence indices were computed for all accessions, and these ranged from 0.44 to 0.83 with a mean index of 0.74. On clustering the accessions via the mixture model, the log likelihood was 593.5 for one group, 730.8 for two groups, 733.9 for three groups and 735.1 for four groups. This indicates that increasing the number of groups from two to three or three to four did not increase the log likelihood value sufficiently to consider three or four groups worthwhile. Therefore, it was decided to use the two-group summary whereby the accessions were allocated to the group to which they had the greater probability of belonging. The cut-off prevalence index value between the groups was 0.691. One of these groups consisted of 58 accessions (14%) with low prevalence indices (subsequently called the 'first group', with average marker prevalence of 0.61±0.06), while the other had 370 accessions (86%) with high prevalence indices (subsequently called the 'second group', with an average marker prevalence of 0.76±0.03). These means and standard deviations were calculated using complete assignment of the accessions into a group rather than using the output of EMMIX, whereby some accessions are partly allocated to more than one group.

Homogeneity analysis

The 428 accessions were subjected to a homogeneity analysis of the data recorded on the 151 molecular markers. The first four components accounted for 43% of the total variation. The homogeneity analysis did not show discrete grouping patterns when the accession points were plotted for the first four components (Fig. 1). These points were densely concentrated except for about 10–15% of the total number of accessions. To combine the results from the clustering and homogeneity analysis, we used a triangle and circle as symbols representing the first and second groups of accessions, respectively (later designated as introgressed and non-introgressed, respectively) in this display of accession

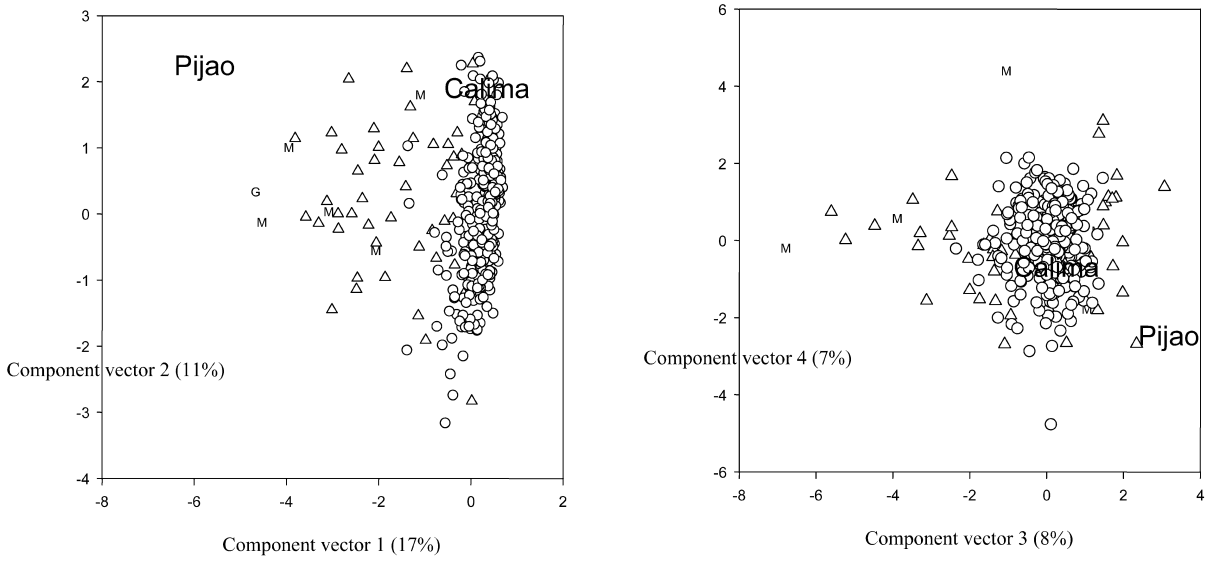


Fig. 1 Plot of the 426 common bean (*Phaseolus vulgaris*) accessions from the Andean gene pool and two checks using the first and second and third and fourth component vectors from the homogeneity analysis of the 428 bean accessions by the 151 molecular-marker array. The percentage of variation accounted for by each component vector is indicated. *Symbols* depict membership at the two-group level from the cluster analysis of the accession by marker prevalence index array: *triangle* Introgression (*G* Guatemala accessions, *M* Mexican accessions), *circle* non-introgression

points in the component space. Accessions were plotted for all combinations of the first four component vectors, but only the plots of component vectors 1 versus 2 and 3 versus 4 were presented here as they adequately described the response patterns. The circle occupied the region in which most accessions were concentrated, while the triangle occupied the outlying regions in the plot (Fig. 1). The Andean check Calima (with a prevalence index of 0.7286) was placed within the densely concentrated area, while the Middle American check ICA-Pijao (with a prevalence index of 0.4403) was placed among those in the outlying regions. The names Calima and Pijao in Fig. 1 are centered on the respective two data points for these accessions (although the name of the former is hard to discern in the second of these pictures).

Geographical distribution of clusters

The geographic origin of the accessions was investigated to see if origin was related to the above patterns. Among the 428 accessions in this study, only 278 accessions had data on longitude and latitude: 41 accessions (out of 58) in the first group and 237 accessions (out of 370) in the second group. On the basis of the geographic information gathered on the 41 accessions in the first group, it was apparent that these 41 accessions were not confined to a particular country of the Andean region but were widely distributed throughout the Andean region of Peru, Ecuador and Colombia. While none of the accessions with data on longitude and latitude in the first group came from Bolivia or Argentina, four out of the 41 were of Middle American origin.

It seems possible that the accessions in the first group had DNA bands which differed from those of the main Andean (or second) group through hybridization or out crossing, possibly with Middle American or North Andean germplasm (i.e. by introgression). Given this hypothesis, it was decided to designate accessions from the first group (denoted by a triangle in Fig. 1) as introgressed with Middle American or North Andean germplasm, while those from the second group (denoted by a circle in Fig. 1) were designated as non-introgressed.

Introgressed accessions

The 58 introgressed accessions together with their geographic origin and prevalence indices are given in Table 1. Six originated in Guatemala and Mexico, while the rest came from the northern and central Andean region. The ICA-Pijao check showed the second lowest prevalence index (0.4403). The one accession from Guatemala and five from Mexico had prevalence indices no greater than 0.5403 and were definitely positioned in the outlying region of Fig. 1. In comparison, the other 12 accessions from Middle America (one from Guatemala, two from Costa Rica and nine from Mexico) had prevalence indices ranging from 0.7383 to 0.7927 and were definitely positioned in the region in which most accessions were concentrated in Fig. 1.

Some introgressed and non-introgressed accessions are shown for primer AK06 (of the ten selected primers) in Fig. 2. With this particular primer, both Andean accessions G16104E and G23484G have bands that were

Table 1 The 58 introgressed accessions found by cluster analysis of the prevalence indices recorded on 426 common bean (*Phaseolus vulgaris*) accessions from the Andean gene pool and two checks

Accessions	Origin	Prevalence indices	Accessions	Origin	Prevalence indices
G 8465	Guatemala	0.4387	G 21242	Colombia	0.6345
Pijao	Colombia	0.4403	G 23604	Peru	0.6350
G 2568	Mexico	0.4609	G 23818A	Peru	0.6362
G 11355	Ecuador	0.4822	G 23814D	Peru	0.6369
G 22033	Mexico	0.4899	G 23782	Peru	0.6418
G 21056	Mexico	0.5218	G 12841	Ecuador	0.6430
G 23564A	Mexico	0.5376	G 23488F	Peru	0.6439
G 2376	Mexico	0.5403	G 8214	Peru	0.6476
G 12858A	Peru	0.5483	G 21227	Colombia	0.6478
G 16345	Argentina	0.5510	G 12403	Ecuador	0.6500
G 23824B	Peru	0.5593	G 12657	Colombia	0.6545
G 23823C	Peru	0.5627	G 7930	Argentina	0.6574
G 19142B	Colombia	0.5740	G 23575G	Peru	0.6599
G 14737	Peru	0.5750	G 7257	Colombia	0.6655
G 17172	Ecuador	0.5774	G 19545	Peru	0.6663
G 8157	Colombia	0.5792	G 17198	Ecuador	0.6699
G 23815	Peru	0.5846	G 9846	Ecuador	0.6715
G 23484G	Peru	0.5858	ABA 2	Colombia	0.6720
G 19368A	Peru	0.5923	G 7437	Colombia	0.6727
G 22287	Peru	0.5963	G 16906	Peru	0.6739
G 7229	Colombia	0.5990	G 12644	Colombia	0.6746
G 23814B	Peru	0.6058	G 14643	Colombia	0.6751
G 12858B	Peru	0.6062	G 17187	Ecuador	0.6776
G 11723	Peru	0.6076	G 12715	Colombia	0.6782
G 16104E	Peru	0.6099	G 23784	Peru	0.6786
G 23829	Peru	0.6185	G 7381	Colombia	0.6887
G 23797A	Peru	0.6203	G 23575D	Peru	0.6891
G 21039	Argentina	0.6216	G 7317	Colombia	0.6905
G 22286	Peru	0.6268			
G 23806A	Peru	0.6300			

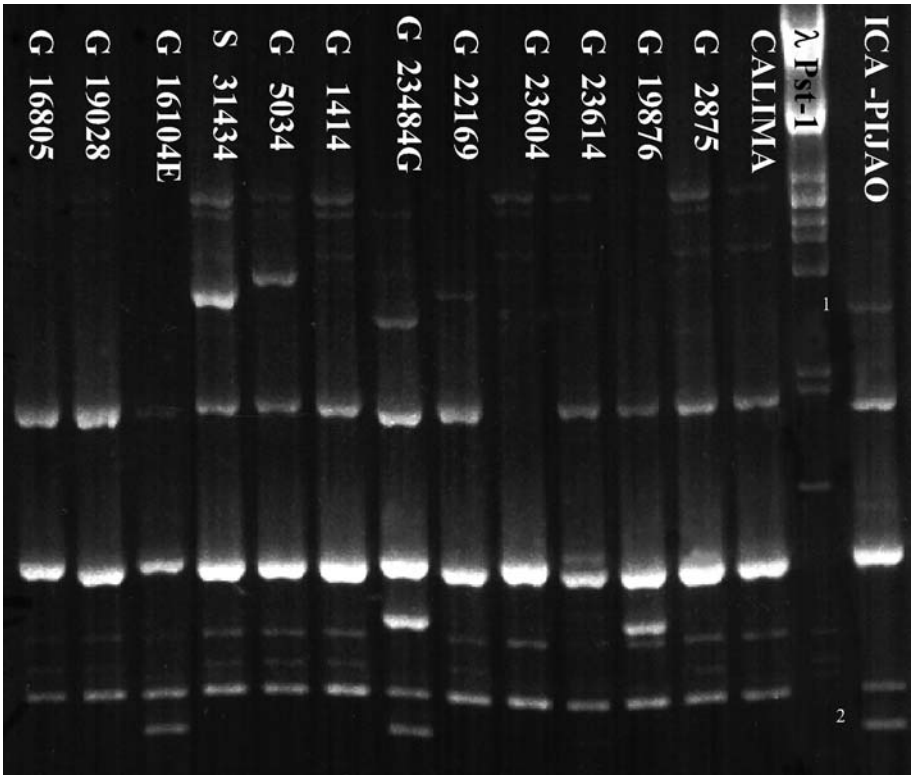


Fig. 2 A RAPD analysis of accessions with Andean-type *phaseolin* from primary centers of origin, indicating presence of Middle American introgression. The Middle American control *(ICA-PIJAO)* has two bands (*1*, *2*) that are not present in the Andean control (*CALIMA*). *λPst-1* is presented as a band size control

Table 2 Mean and standard deviation (SD) of disease, pest and nutritional attributes and the percentage of resistant accessions of both introgressed and non-introgressed groups of 426 common bean accessions from the Andean gene pool and the *p* values for the *t*-test between the group means

Attributes[a]	Introgressed group (G1)		Non-introgressed group (G2)		G1 versus G2 *p* values
	Mean±SD	Percentage of resistant accessions	Mean±SD	Percentage of resistant accessions	
ALS-a	3.56±2.01	57.9	5.42±1.90	19.8	0.000**
ANT-a	4.07±2.60	45.6	5.35±2.91	31.4	0.002**
ALS-m	6.30±1.57	10.7	5.47±1.76	20.1	0.001**
ANT-m	2.71±2.57	73.2	1.89±2.00	83.5	0.000**
CBB	6.16±1.15	3.5	6.14±1.17	3.0	0.893
Empoasca	8.25±0.43	0	8.18±0.43	0	0.262
DF	65.1±12.0		66.6±12.1		0.378
DM	27.8±13.9		130.8±15.0		0.157
SWt	43.8±17.1		57.7±16.0		0.000**
PC (%)	22.7±2.5		20.8±2.1		0.000**
Ca (%)	0.142±0.044		0.137±0.041		0.410
P (%)	0.365±0.082		0.337±0.066		0.007**
S (%)	0.222±0.025		0.204±0.021		0.000**
Fe (mg/kg)	59.0± 9.9		55.7±7.9		0.008**
Zn (mg/kg)	35.0±5.3		32.2±4.8		0.000**

** Significant at $p \leq 0.01$

[a] ALS, Angular leaf spot; ANT, anthracnose (a, Andean; m, Middle American); CBB, common bacterial blight; DF, days to flowering; DM, days to maturity; SWt, seed weight in grams per 100 seed; PC, protein concentration (%); Ca, calcium; P, phosphorus; S, sulfur; Fe, iron; Zn, zinc

present in the Middle American control (ICA-Pijao) but not in the Andean control (Calima).

Effects of introgression on quantitative traits

The group means and standard deviations for morphological (DF, DM and SWt) and nutrient elements (protein, Ca, P, S, Fe and Zn) and percentage of resistant accessions for the disease and pest attributes (ALS-a, ANT-a, ALS-m, ANT-m, CBB and *Empoasca*) are shown in Table 2. Independent sample *t*-tests showed that for ten attributes (ALS-a, ANT-a, ALS-m, ANT-m, SWt, protein, Fe, P, S and Zn), there were significant differences between the means of the introgressed and non-introgressed accessions at the 1% level.

The attribute values of the accessions for introgressed and non-introgressed groups were plotted against the marker prevalence indices (Fig. 3) to demonstrate the variability of accession responses. These plots supported the findings that the introgressed group (with lower prevalence indices) was associated with higher resistance to ALS-a (scores ≤ 3) and ANT-a but exhibited more susceptibility to ALS-m and ANT-m (scores >3).

Discussion

The position of the outliers tending towards the Middle American check (Fig. 1) suggests that the former were similar to the Middle American germplasm and different from most of the Andean germplasm. Although most of the outliers originated in the Andean region, it appears that by natural crossing between members of different gene pools they have acquired DNA bands similar to the Middle American germplasm. It is likely that this has occurred from landraces that were introduced to the Andean region, either in the pre-Colombian period or later.

On the other hand, six of the introgressed accessions originated in Guatemala and Mexico (Table 1), while another 12 accessions from Middle America showed no significant effects of introgression. When compared to the proportion of introgressed accessions in the Andean region (52 out of 408, or 13%), the Middle American accessions with Andean phaseolin actually showed a higher degree of introgression (six out of 18, or 33%). These Andean types are a small minority in Middle America and are surrounded by landraces that are genetically Middle American. Thus, there is a higher probability that any outcrossing of these Andean types would involve Middle American germplasm than would occur in the Andean region. This might explain the surprisingly high frequency (33%) of introgressed Andean types in the Middle American region compared to the Andean region.

Most Andean accessions formed a tight cluster in the homogeneity analysis of RAPD data (Fig. 1), thus confirming previous results with AFLP that suggested a narrow genetic base of Andean germplasm (Beebe et al. 2001). This previous study with AFLP dealt with a smaller set of accessions (of 121), and only 16 introgressed accessions were identified. Thus, it was not possible to compare the effects of introgression on phenotypic traits. In the present study, 58 accessions were identified as introgressed, and this permitted a comparison of phenotypic traits of introgressed and non-introgressed accessions. The introgressed group was

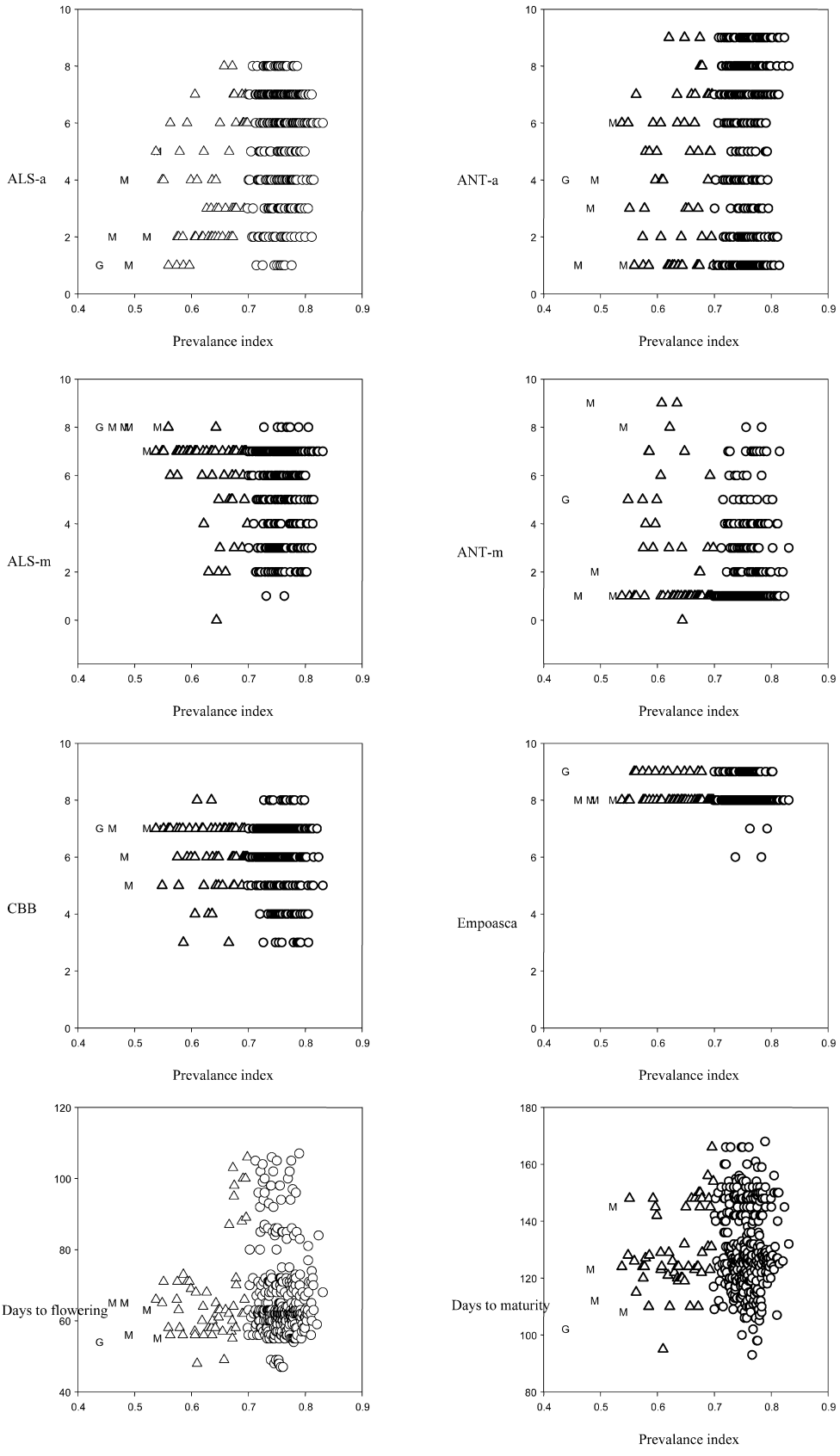
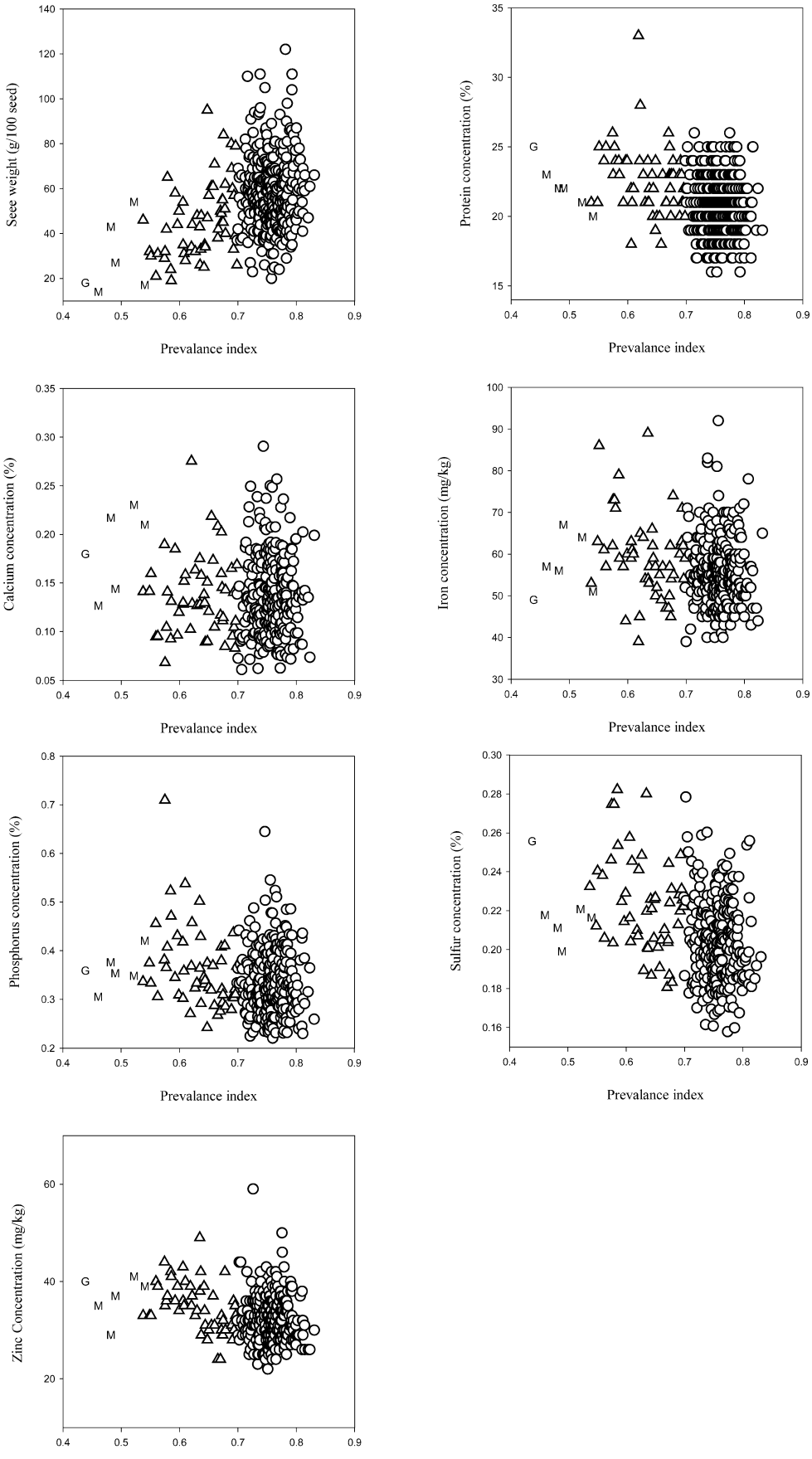


Fig. 3 Plot of attribute value against prevalence index for the 426 common bean accessions from the Andean gene pool and two checks. *Triangle* Introgression (*G* Guatemala, *M* Mexican accessions), *circle* non-introgression

Fig. 3 (continued)



associated with lower seed weight, higher resistance for ALS-a and ANT-a, higher protein, Fe, S, P and Zn concentrations and a lower resistance for ALS-m and ANT-m than the non-introgressed group. Therefore, these traits were associated with lower marker prevalence indices, or the presence of rare bands. All of these traits are also typical of the Middle American gene pool. For example, in the case of resistance to angular leaf spot and anthracnose, evidence exists that these two pathogens have co-evolved with their host and are adapted to the respective gene pools (Pastor-Corrales et al. 1998). Thus, susceptibility to Middle American isolates suggests that the DNA fragments that contributed to low marker prevalence originated in the Middle American gene pool. Resistance to Andean isolates suggests the same. The lower prevalence index that was associated with low seed weight is consistent with the hypothesis that the rare bands come from beans of the Middle American gene pool, since these possessed smaller seed than the Andean types (Singh et al. 1991a). The lower prevalence index associated with higher protein, P, S, Fe and Zn concentrations indicates that rare markers were associated with higher nutrient elements. The Middle American gene pool presented higher values of these elements than the Andean gene pool (Islam et al. 2002b). The significantly higher Fe level in the introgressed group was similar to the North Andean group (unique for the presence of C and H phaseolin) reported by Islam et al. (2002a).

The implication of the foregoing is that introgression among gene pools has started to influence the evolution of common bean landraces of the Andean gene pool. Middle American genes are influencing a range of traits from disease resistance to seed morphology. This is an important fact to keep in mind as breeders and geneticists search for unique genes or gene combinations among landraces. For example, if a landrace presents an unusual degree of resistance to a local pathogen, it might result from just such an introgression. With respect to mineral concentration, accession G21242, which displayed evidence of introgression, proved to have usually high seed iron concentration (unpublished data). Introgression may explain some other unique cases of exceptional germplasm in the Andean gene pool.

The accessions in the introgressed group were matched with the groups reported by Islam et al. (2002a). Of the 38 accessions in the North Andean gene pool in that study, only 18 were included in the 426 Andean accessions subjected to RAPD analysis (a proportion of 0.042). The proportion of North Andean accessions in the introgressed group (three out of 58, or 0.052) was close to that expected by chance. This result led us to conclude that the rare DNA bands associated with lower marker prevalence indices do not reflect an influence of North Andean beans and were probably derived from the Middle American gene pool.

Middle American beans were cultivated in the Andean zone prior to the colonial period (Towle 1961; Kaplan and Kaplan 1988). Archaeological evidence suggests that trade between Central America and the Andean zone was active in pre-Colombian times, and this could have established Middle American beans in the Andes (Gepts 1993). The introgression detected in the present study therefore could be very ancient.

Introgression among gene pools is likely to produce off-types for seed colour, shape and size and other attributes. Such types will not be acceptable in production systems oriented toward markets where commercial criteria limit the types of seed that can be marketed. Off-types will be maintained principally in subsistence systems for home or local consumption. Systems that are truly subsistence and isolated from market forces are disappearing rapidly, and it is to be expected that introgression may not continue in the future.

Introgression among the Middle American and Andean gene pools has been a difficult challenge for bean breeders. Intergene pool progeny are often poor agronomically, and directed, concentrated efforts have been necessary to make effective use of crosses among Andean and Middle American genotypes (Beaver and Kelly 1994). In light of this, it is striking that some introgressed genotypes have been sufficiently productive to have survived in farmers' systems, possibly due to more effective disease resistance. Some of these may represent unique genetic recombination events that could be of utility to breeders seeking to improve common bean.

Acknowledgements This research was financially supported by a grant from the Australian Centre for International Agricultural Research (ACIAR), Canberra, ACT. The experiments and analyses reported in this paper comply with the current laws of both Colombia and Australia. Constructive comments from a reviewer and the Editor were very helpful.

References

Balardin RS, Kelly JD (1998) Interaction between *Colletotrichum lindemuthianum* races and gene pool diversity in *Phaseolus vulgaris*. J Am Soc Hortic Sci 123:1038–1047

Basford KE, McLachlan GJ (1985) The mixture method of clustering applied to three-way data. J Class 2:109–125

Beaver JS, Kelly JD (1994) Comparison of selection methods for dry bean populations derived from crosses between gene pools. Crop Sci 34:34–37

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

Beebe SE, Toro Ch O, Gonzalez AV, Chacón MI, Debouck DG (1997) Wild-weedy 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, Skroch PW, Tohme J, Duque MC, Pedraza F, Nienhuis J (2000) Structure of genetic diversity among common bean landraces of Middle American origin based on correspondence analysis of RAPD. Crop Sci 40:264–273

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

Beer SC, Siripoonwiwat W, O'Donoughue LS, Souza E, Matthews D, Sorrells ME (1997) Association between molecular markers and quantitative traits in an oat germplasm pool: can we infer linkages? http://www.ncgr.org/ag/jag/papers97/paper197/jqte1999-01.html

Debouck DG, Araya Villalobos R, Ocampo Sánchez RA, González UWG (1989) Collecting *Phaseolus* in Costa Rica. FAO/IBPGR Plant Genet Resour Newsl 78/79:44-46

Geffroy V, Sicard D, de Oliveira JCF, Sevgnac M, Cohen S, Gepts P, Neema C, Langin T, Dron M (1999) Identification of an ancestral resistance gene cluster involved in the co-evolution process between *Phaseolus vulgaris* and its fungal pathogen *Colletotrichum lindemuthianum*. Mol Plant Microb Interact 12:774–784

Gepts P (1988a) A Middle American and an Andean common bean gene pool. In: Gepts P (ed) Genetic resources of *Phaseolus* beans. Kluwer, Dordrecht, pp 375–390

Gepts P (1988b) Phaseolin as an evolutionary marker. In: Gepts P (ed) Genetic resources of *Phaseolus* beans. Kluwer, Dordrecht, pp 215–241

Gepts P (1993) The use of molecular and biochemical markers in crop evolution studies. In: Hecht MK (ed) Evolutionary biology. Plenum, New York, pp 51–94

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

Gepts P, Debouck DG (1991) Origin domestication and evolution of the common bean (*Phaseolus vulgaris* L.). In: van Schoonhoven A, Voysest O (eds) Common beans: research for crop improvement. Commonwealth Agricultural Bureau, Wallingford, pp 7–53

Guzmán P, Gilbertson RL, Nodari R, Johnson WC, Temple SR, Mandela D, Mkandawire ABC, Gepts P (1995) Characterization of variability in the fungus *Phaseoisariopsis griseola* suggests coevolution with the common bean (*Phaseolus vulgaris*). Phytopathology 85:600–607

Islam FMA, Basford KE, Redden RJ, Gonzalez AV, Kroonenberg PM, Beebe SE (2002a) Genetic variability in cultivated common bean beyond the two major gene pools. Genet Resour Crop Evol 49:271–283

Islam FMA, Basford KE, Jara C, Redden RJ, Beebe SE (2002b) Seed compositional and disease resistance differences among gene pools in cultivated common bean. Genet Resour Crop Evol 49:285–293

Islam FMA, Basford KE, Redden RJ, Jara C, Beebe SE (2002c) Patterns of resistance to angular leaf spot, anthracnose and common bacterial blight in common bean germplasm. Aust J Exp Agric 42:481–490

Jones PG, Gladkov A (1999) Flora map. In: Jones AL (ed) Version 1.1. Manual, copyright CIAT 1998–1999. CIAT, Colombia

Jones PG, Galway N, Beebe SE, Tohme J (1997) The use of geographical information systems in biodiversity exploration and conservation. Biodivers Conserv 6:947–958

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

Kaplan L, Kaplan LN (1988) *Phaseolus* in archaeology. In: Gepts P (ed) Genetic resources of *Phaseolus* beans. Kluwer, Dordrecht, pp 125–142

Khairallah MM, Adams MW, Sears BB (1990) Mitochondrial DNA polymorphisms of Malawian bean lines: further evidence for two major gene pools. Theor Appl Genet 80:753–761

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

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

Manly BFJ (1994) Multivariate statistical methods, 2nd edn. Chapman and Hall, New York

Paredes OM, Gepts P (1995) Extensive introgression of Middle American germplasm into Chilean common bean cultivars. Genet Resour Crop Evol 42:29–41

Pastor-Corrales MA (1991) Estandarización de variedades diferenciales y de designación de razas de *Colleototrichum lindemuthianum*. Phytopathology 81:694

Pastor-Corrales MA, Jara C, Singh SP (1998) Pathogenic variation in sources of and breeding for resistance to *Phaeoisariopsis griseola* causing angular leaf spot in common bean. Euphytica 103:161–171

Peel D, McLachlan GJ (1999) User's guide to EMMIX, version 10 http://www.maths.uq.edu.au/~gjm/emmix/guide.html

Schoonhoven AV, Pastor-Corrales MA (1987) Standard system for the evaluation of bean germplasm. Centro Internacional de Agricultura Tropical, Cali, Colombia

Singh SP (1995) Selection for seed yield in Middle American versus Andean × Middle American interracial common bean populations. Plant Breed 114:269–271

Singh SP, Gutiérrez JA (1990) Effect of plant density on selection for seed yield in two population types of *Phaseolus vulgaris* L. Euphytica 51:173–178

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

Singh SP, Gutiérrez JA, Molina A, Urrea C, Gepts P (1991b) Genetic diversity in cultivated common bean II Marker-based analysis of morphological and agronomic attributes. Crop Sci 31:23–29

Singh SP, Terán H, Muñoz CG, Takegami JC (1999) Two cycles of recurrent selection for seed yield in common bean. Crop Sci 39:391–397

Skroch PW, Nienhuis J, Beebe S, Tohme J, Pedraza F (1998) Comparison of Mexican common bean (*Phaseolus vulgaris* L.) core and reserve germplasm collections. Crop Sci 38:488–496

Snedecor GW, Cochran WG (1989) Statistical methods, 8th edn. Iowa State University Press, Ames

SPSS for Windows (1999) Release 901 (24 Feb 1999) Copyright SPSS 1989–1999

Tohme J, Jones P, Beebe SE, Iwanaga M (1995) The combined use of agroecological and characterisation data to establish the CIAT *Phaseolus vulgaris* core collection. In: Hodgkin T, Brown AHD, van Hintum ThJL, Morales EAV (eds) Core collections of plant genetic resources. International Plant Genetic Resources Institute (IPGRI). Wiley-Sayce Publ, London, pp 95–107

Towle MA (1961) The ethnobotany of pre-Columbian Peru. Aldine, Chicago

White J, Kornegay J, Castillo J, Molano CH, Cajiao C, Tejada G (1992) Effect of growth habit on yield of large-seeded bush cultivars of common bean. Field Crops Res 29:151–161