Theor Appl Genet (2010) 120:1635–1651
DOI 10.1007/s00122-010-1282-0

ORIGINAL PAPER

Integration of genome and phenotypic scanning gives evidence of genetic structure in Mesoamerican common bean (*Phaseolus vulgaris* L.) landraces from the southwest of Europe

M. Santalla · A. M. De Ron · M. De La Fuente

Received: 9 June 2009 / Accepted: 24 January 2010 / Published online: 9 February 2010
© Springer-Verlag 2010

Abstract Southwestern Europe has been considered as a secondary centre of genetic diversity for the common bean. The dispersal of domesticated materials from their centres of origin provides an experimental system that reveals how human selection during cultivation and adaptation to novel environments affects the genetic composition. In this paper, our goal was to elucidate how distinct events could modify the structure and level of genetic diversity in the common bean. The genome-wide genetic composition was analysed at 42 microsatellite loci in individuals of 22 landraces of domesticated common bean from the Mesoamerican gene pool. The accessions were also characterised for phaseolin seed protein and for nine allozyme polymorphisms and phenotypic traits. One of this study's important findings was the complementary information obtained from all the polymorphisms examined. Most of the markers found to be potentially under the influence of selection were located in the proximity of previously mapped genes and quantitative trait loci (QTLs) related to important agronomic traits, which indicates that population genomics approaches are very efficient in detecting QTLs. As it was revealed by outlier simple sequence repeats, loci analysis with STRUCTURE software and multivariate analysis of phenotypic data, the landraces were grouped into three clusters according to seed size and shape, vegetative growth habit and genetic resistance. A total of 151 alleles were detected with an average of 4 alleles per locus and an average polymorphism information content of 0.31. Using a model-based approach, on the basis of neutral markers implemented in the software STRUCTURE, three clusters were inferred, which were in good agreement with multivariate analysis. Geographic and genetic distances were congruent with the exception of a few putative hybrids identified in this study, suggesting a predominant effect of isolation by distance. Genomic scans using both markers linked to genes affected by selection (outlier) and neutral markers showed advantages relative to other approaches, since they help to create a more complete picture of how adaptation to environmental conditions has sculpted the common bean genomes in southern Europe. The use of outlier loci also gives a clue about what selective forces gave rise to the actual phenotypes of the analysed landraces.

Communicated by R. Varshney.

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

Introduction

Following the initial domestication phase, the common bean (*Phaseolus vulgaris* L.) spread between Mesoamerica and South America and, after the European exploration of the Americas, to Europe and Africa (Gepts and Bliss 1988; Gepts et al. 1986; Gepts 1988) where it was cultivated under diverse environments and farmer preferences. Thanks to its successful cultivation in Spain and Portugal, the common bean found a secondary centre for diversification in the Iberian Peninsula (Santalla et al. 2002; Rodiño et al. 2006). The common bean consists of two major gene pools, Mesoamerican from Central America and Mexico, and Andean from the Andes Mountains of South America (Gepts et al. 1986; Koenig and Gepts 1989; Singh et al. 1991a–c; Becerra Velasquez and Gepts 1994; Tohme et al. 1996; Beebe et al. 2000, 2001), as determined by

morphological attributes (Singh et al. 1991a, b), seed proteins (Gepts 1988), allozymes (Koenig and Gepts 1989) and DNA polymorphisms (Khairallah et al. 1992). The existence of the Andean and Mesoamerican gene pools in the common bean and the multiple domestications associated with them, are a unique situation among crops, rice being an exception (Vitte et al. 2004; Londo et al. 2006). The existence of these two gene pools raises a number of questions such as the origin, dispersion and relationships between these two gene pools, the qualitative and quantitative differences in genetic diversity between them, the respective levels of linkage disequilibrium, and the extent to which different loci have been the subject of selection during and after the two major domestication events in the species.

Bean gene pools appear to have co-evolved with local pathogen landraces, since the accessions in one gene pool are often susceptible to pathogen races from the same region but relatively more resistant to those from other regions (Guzmán et al. 1995; Pastor-Corrales et al. 1998; Balardin and Kelly 1998; Geffroy et al. 1999; Islam et al. 2002). The Mesoamerican gene pool is composed of races Durango, Jalisco, Mesoamerica and Guatemala (Beebe et al. 2000; Singh et al. 1991a; Chacón et al., 2005; Díaz and Blair 2006; Blair et al. 2007). Races Mesoamerica (tropical black and red Central American beans; carioca beans in Brazil) and Durango (pinto, bayos, and small red and great northern beans) are probably the most widely grown. The race Jalisco is overlapped with the race Durango in terms of seed colour, growth habit and geographical origin. In molecular studies, they have proved to be similar as well (Singh et al. 1991c; Beebe et al. 2000; McClean et al. 2004; Rosales-Serna et al. 2005; Chacón et al. 2005). Singh and Gutiérrez (1990) reported yield gains in inter-racial crosses within the Mesoamerican gene pool; however, there has been a slower progress in the improvement of the Andean types (Kornegay et al. 1992; White et al. 1992). One strategy for the improvement of the Andean types can be achieved by crossing them with the Mesoamerican beans, although these crosses are often not productive (Singh 1995). Some extra efforts, such as a recurrent selection, are necessary to obtain a useful progeny from the inter-gene 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.

Microsatellites or simple sequence repeats (SSRs) have been used with great success to effectively and rapidly estimate the genetic variability within and between landraces, and among accessions of cultivated or wild species (Mitchell et al. 1997). Microsatellite markers are more polymorphic (Blair et al. 2006a) than the other markers used earlier to characterise genetic diversity such as the phaseolin seed protein (Gepts et al. 1986), allozymes (Koenig and Gepts 1989; Singh et al. 1991c), RFLP (Becerra Velasquez and Gepts 1994), and RAPD (Freyre et al. 1996). They are also more widely distributed in the bean genome (Freyre et al. 1998; Blair et al. 2003). In the common bean, around 400 microsatellite markers have been developed and mapped (Yu et al. 2000; Gaitán-Sólis et al. 2002; Blair et al. 2003; Masi et al. 2009; Yaish and Pérez de la Vega 2003; Guerra-Sanz 2004; Caixeta et al. 2005; Buso et al. 2006). However, so far, germplasm studies with microsatellites in the common bean have only been performed in a small number of landraces or breeding lines or they have focussed on certain geographic regions (Metais et al. 2002; Blair et al. 2006a; Díaz and Blair 2006).

One of the single most important and generalised features of plant domestication is the reduction in genetic diversity, not only during the initial domestication phase itself but also subsequently during dispersion from centres of domestication (Gepts 2004). This reduction is caused by both stochastic events (for example, a bottleneck and genetic drift due to a reduction in the population size) and selection (for example, for adaptation to a novel cultivated environment) (Vigouroux et al. 2002). The reduction of genetic diversity has also been more drastic in autogamous species that have restricted effective recombination as compared with allogamous species (Jarvis and Hodgkin 1999). Gene flow between both gene pools appears to be relatively common in the Andean (Debouck et al. 1989; Paredes and Gepts 1995; Beebe et al. 1997; Chacón et al. 2005) and European zones (Santalla et al. 2002; Sicard et al. 2005; Piergiovanni et al. 2006; Rodiño et al. 2006; Sánchez et al. 2008). Inferred evidence of hybridisation due to the presence of morphological intermediate plants is relatively weak, since it may result from either a phenotypic plasticity or a convergent evolution rather than from a gene flow. The presence of crop-specific alleles in morphologically intermediate landraces can help to provide strong evidence for a history of hybridisation. Molecular analyses in conjunction with phenotypic studies of germplasm are recommended because they provide complementary information and increase the resolving power of the genetic diversity (Singh et al. 1991b). Recent approaches—that have been referred to as ‘population genomics’—have a great potential for the detection of QTLs for adaptive traits, with better resolution than the conventional QTL analysis and without the need for a priori knowledge of the phenotypic trait or the candidate gene that may be responsible for the adaptive response (Papa et al. 2005; Rossi et al. 2009).

In the research reported here, a collection of Mesoamerican common bean landraces from Spain and Portugal and

suitable control genotypes was characterised through morphological descriptors, agronomic and disease resistance traits, proteins and SSR markers. The integration of these data could yield significant additional insights into the organisation of the common bean's population structure and help to disclose the role of genetic adaptation to novel environments in changing genetic diversity. Moreover, we have also examined the role of neutral and outlier loci to go into detail of their implication in bean adaptation.

Materials and methods

Plant material

A total of 22 common bean landraces were chosen based on previous designations (Santalla et al. 2002; Rodiño et al. 2006) and the Great Northern seed pattern of each genotype was known to be typical of the Mesoamerican gene pool (Table 1). These 22 landraces were collected from the most important dry bean production areas in the north of Spain and Portugal. The original accessions were grown to evaluate the presence of morphologically heterogeneous stocks. Of the 22 accessions of the cultivated common bean, 15 gave rise to heterogeneous progeny; 154 individuals—about 5–10 individuals per landrace (depending on the heterogeneity showed by each landrace)—were randomly chosen to represent the whole variation. These individuals were regrown to establish uniform accessions. The Mesoamerican cultivars: Alert, Beryl, Matterhorn, USGN-5, USWA-12, USWA-13, Weighing and Almonga were considered as references of the Great Northern type. Three additional genotypes, Calima and Michigan Dark Red Kidney (MDRK), were used as the Andean check cultivars and the ICA Pijao was used as the Mesoamerican check cultivar. A field trial was established with uniform accessions in two locations in northwest Spain: Salcedo (42°24′N, 8°38′W, 60 masl) and Soutomaior (42°19′N, 8°33′W, 290 masl), in two growing seasons (2003 and 2004). The accessions were arranged in a lattice design with two replications and 15 plants per elementary experimental unit. Plants were grown in open fields with standard agronomic practices.

Genomic DNA extraction and genotyping microsatellite

Genomic DNA was extracted from young trifoliate leaves of 10-day-old greenhouse-grown plants from a bulk of eight individuals per accession using the CTAB method (Afanador et al. 1993). Sixty-two microsatellite markers from all 11 linkage groups were chosen based on their

Table 1 Domesticated accessions of the common bean from Spain and Portugal used for SSR analysis

Landrace no.	Common name	Seed weight	Country	Department	County	Altitude (masl)	Latitude (decimal)	Longitude (decimal)
PHA-0028	Garabanzo	44.59	Spain	Pontevedra	Tui	26	42.05	−8.64
PHA-0055	Feijao manteiga	44.75	Portugal	Bragança	Macedo de Cavaleiros	648	41.54	−6.96
PHA-0071	Feijao manteiga	50.39	Portugal	Vila Real	Valpaços	484	41.61	−7.31
PHA-0098	Feijao manteiga	60.15	Portugal	Bragança	Vila Flor	545	41.31	−7.15
PHA-0103	Feijao manteiga	59.30	Portugal	Bragança	Mirandela	249	41.49	−7.18
PHA-0181	Faba do caldo	55.11	Spain	Lugo	Viveiro	118	43.66	−7.59
PHA-0229	Garabanzo	40.64	Spain	Pontevedra	Campo Lameiro	225	42.54	−8.54
PHA-0245	Faba do caldo	56.40	Spain	Pontevedra	A Estrada	271	42.69	−8.49
PHA-0294	Garabanzo	46.54	Spain	Orense	Nogueira de Ramuin	619	42.41	−7.73
PHA-0397	Chichos	47.19	Spain	Asturias	Pravia	20	43.49	−6.11
PHA-0399	Faba pancha	75.38	Spain	Asturias	Pravia	20	43.49	−6.11
PHA-0418	Garabanzo	48.46	Spain	Pontevedra	Mos	211	42.19	−8.65
PHA-0419	Habichuela	64.01	Spain	Pontevedra	Vilanova de Arousa	2	42.56	−8.83
PHA-0593	Ganxet	40.43	Spain	Leon	Villamañan	764	42.32	−5.58
PHA-0609	Planchada	64.46	Spain	Leon	La Bañeza	829	42.61	−5.57
PHA-0610	Plancheta	42.50	Spain	Leon	La Bañeza	829	42.61	−5.57
PHA-0623	Ganxet	43.09	Spain	Cataluña	Cardedeu	195	41.38	2.21
PHA-1014	Faba do caldo	53.64	Spain	Lugo	Alfoz	93	43.53	−7.41
PHA-1057	Feijao manteiga	52.46	Portugal	Vila Real	Borbela	520	41.33	−7.75
PHA-1058	Feijao manteiga	45.18	Portugal	Vila Real	Borbela	520	41.33	−7.75
PHA-1061	Feijao manteiga	46.20	Portugal	Viseu	Moimenta da Beira	690	40.98	−7.62
PHA-1068	Plancheta	45.28	Spain	Leon	Villamañan	764	42.32	−5.58

dispersed map location (Yu et al. 2000; Gaitán-Sólis et al. 2002; Blair et al. 2003, 2006a). With the exception of marker pairs BM53 and BM200 (linkage group B01), BMd2 and BMd7 (linkage group B02), and BMd9 and BMd16 (linkage group B04), which were each closely linked, all other pairs were spaced by 1 cM or more apart. The genetic linkage map location, repeat motif, and primer sequences can be found in the original publications (BMd: Blair et al. 2003, 2006a; Pv: Yu et al. 2000; BM: Gaitán-Sólis et al. 2002). PCR reaction mixtures contained approximately 50 ng of total genomic DNA, 1 U of Taq polymerase (BioTaq™ DNA polymerase, Biolife), 0.2 mM of each dNTP, 5–7.5 pmol (depending on the primer pair) of each primers, 1× PCR buffer and 1.5–25 mM $MgCl_2$ in a 25 µl total reaction volume. PCR cycles, performed on a GeneAmp PCR System 9700 thermocycler (Applied Biosystems, Foster City, CA, USA), consisted of 5 min at 94°C and 30 cycles of 1 min at 94°C, 1 min at 47–68°C (depending on the set of primers used) and 1 min at 72°C, and a final extension of 30 min at 72°C. The amplified fragments were multiplexed depending on their size variation and analysed in an ABI PRISM 3130xl genetic analyser with a 600LIZ size standard internal weight marker (35 fragments from 20 to 600 bp) (Applied Biosystems Inc.) and a Performance Optimised Polymer 7 (Applied Biosystems, Inc.). Genotypes for each marker were determined using the Genescan 3.7 and GeneMapper software (Applied Biosystems Inc.).

Protein extraction and phaseolin and allozyme pattern

Total proteins were extracted from 0.1 g of an individual peeled seed from five individuals per accession and the phaseolin pattern of each genotype was analysed using SDS-PAGE electrophoresis as described by Brown et al. (1981) and Santalla et al. (2002). The primary leaf and root apex tissues (depending on the enzymes assayed) of 15-day-old seedlings from six individuals per accession were used for allozyme analysis following the method of Santalla et al. (2002). Six polymorphic enzyme systems were assayed: malic enzyme (*Me*; E.C.1.1.1.40), shikimate dehydrogenase (*Skdh*, E.C.1.1.1.25), ribulose bisphosphate carboxylase (*Rbcs*, E.C.4.1.1.39), peroxidase (*Prx*, E.C.1.11.1.7), malate dehydrogenase (*Mdh*, E.C.1.1.1.37) and diaphorase (*Diap*, E.C.1.6.99). The *Mdh* and *Diap* enzyme systems had two independent loci each. Loci were labelled sequentially with those migrating closest to the anodal end being designated as number 1 (Koenig and Gepts 1989). The most common allele was designated as 100 and all other alleles were measured in millimetres from the standard. ICA Pijao and MDRK cultivars were included as standards. ICA Pijao and MDRK cultivars have respectively the following genotypes at polymorphic enzyme loci: $Rbcs^{100}$, $Skdh^{103}$, Prx^{98}, Me^{100}, $Mdh\text{-}1^{100}$, $Mdh\text{-}2^{100}$, $Diap\text{-}1^{95}$ and $Diap\text{-}2^{105}$, and $Rbcs^{98}$, $Skdh^{100}$, Prx^{98}, Me^{98}, $Mdh\text{-}1^{103}$, $Mdh\text{-}2^{102}$, $Diap\text{-}1^{100}$ and $Diap\text{-}2^{100}$, respectively.

Phenotypic diversity

On a single plant basis, 11 morphological traits were scored. The traits recorded were: (1) leaf width in centimetres of three centre trifoliolate leaves, (2) leaf length in centimetres of three centre trifoliolate leaves, (3) leaf shape recorded as chordate, ovate, rhombohedric or hastate following the classification developed by Singh et al. (1991a), (4) growth habit classified by CIAT (1987) 1–4 scale, (5) length of the fifth internode on the main stem measured in centimetres, (6) number of nodes on the main stem to first flower, (7) bracteole shape classified as cordate, ovate, lanceolate or triangular, (8) bracteole size classified as small, medium or large, (9) dry seed length in millimetres, (10) dry seed height in millimetres and (11) dry seed width in millimetres. Most of the morphological traits were measured at the R1 (initiation of flowering) stage of development (CIAT 1987). Two phenological traits, days to first flowering and days to maturity, were recorded on the basis of the average of the entire plot (CIAT 1987). Agronomical traits recorded were pods per plant, seeds per pod, seed yield (kg ha^{-1} at 140 g kg^{-1} of moisture) and 100-seed weight in grams.

Reaction to anthracnose (ANT), caused by the fungus *Colletotrichum lindemuthianum*, was evaluated using Andean (7, 23, 39, 55 and 102 races) and Mesoamerican (17, 73, 448 and 1,545 races) pathogen isolates in independent inoculations (Pastor-Corrales 1991). The common bacterial blight (CBB) incited by *Xanthomonas axonopodis* pv. *phaseoli* (Smith) was evaluated on leaves. Nurseries were inoculated with a mixture of 241, 260, 326 and 410 strains after planting. Data on these two diseases were taken on a 1–9 scale (1, immune; and 9, very susceptible) according to van Schoonhoven and Pastor-Corrales (1987). The presence of the dominant *I* and recessive *bc* genes for resistance to the bean common mosaic virus (BCMV/BCMNV) was verified in glasshouse inoculations on 7- to 10-day-old seedlings with the NL3 and US6 strains. Reaction to the halo bacterial blight (HB), caused by *Pseudomonas syringae* pv. *phaseolicola*, was studied on leaves and pods. The two primary leaves of the seedlings and at least two immature pods per plant were inoculated with an isolate of race 3 (isolate 2425). Symptoms were observed 10 days after the inoculation in both leaves and pods. The susceptibility level on leaves and pods was evaluated on a 5-level scale according to Innes et al. (1984), as follows: (1) red-brown necrotic reaction in the area of maximum inoculation (highly resistant), (2) red-brown necrotic reaction with a trace of water soaking (resistant), (3) some necrosis but more extensive water-soaking largely confined

to the area of maximum inoculation (slightly susceptible), (4) small water-soaked lesions (<1 mm diameter) distributed over the leaf surface (susceptible) and (5) larger water-soaked lesions (>1 mm diameter) distributed over the leaf surface (highly susceptible).

Detection of loci under selection

Two different approaches were used to identify microsatellite loci that may have been under selection. Both of them are based on the straightforward principle that genetic differentiation among populations is expected to be different for loci under a divergent selection than for the rest of the genome. Both approaches use computer simulations under a defined evolutionary scenario, and loci lying outside the neutral distribution are detected as outliers. The Beaumont and Nichols (1996) approach, implemented in LOSITAN (Antao et al. 2008), use computer simulations to detect loci where the genetic diversity within (heterozygosity) and between accessions (F_{ST}) do not conform to the prediction of a simple infinite or finite island model obtained by coalescent simulations. Loci with unusually high or low F_{ST} values compared to neutral expectations are considered to be potentially under a divergent selection. The observed F_{ST} estimates are compared with the simulation output (the simulated distribution of F_{ST}) to identify potential outliers. This procedure lowers the bias on the estimation of the mean neutral F_{ST} by removing the most extreme loci from the estimation. Ten thousand realisations were used at each step. Pairwise analyses were subsequently performed for each possible accession combination (10 single analyses) to check for the presence of loci with atypical behaviour due to chance alone, which are not expected to exhibit parallel trends in several comparisons (Campbell and Bernatchez 2004; Storz 2005; Vasemägi et al. 2005).

The second programme used to detect outlier loci was DetSel, which relies on a model in which the common ancestor population splits into two populations that later diverge only by random drift (Vitalis et al. 2001, 2003). Population-specific parameters are defined as functions of identity probabilities for pairs of genes taken within or between populations. For our analysis, a null distribution was obtained based on a range of nuisance parameters [mutation rate (μ) varying from 10^{-2} to 10^{-4}, ancestral accession size before the bottleneck (N_e) from 10^3 to 10^4, ancestral accession size during the bottleneck (N_0) from 165 and ratio of divergence time (t) from 80 to 50] which generated a similar number of alleles as in the observed data set [the number of generations during the bottleneck (τ_0) was equal to 100]. Fifty thousand realisations were performed for each possible pair of single populations. The significance level was set at 0.99 in order to obtain the most conservative results. Loci that were monomorphic in only one of the two compared accessions were systematically discarded from the simulations because the confidence envelope appeared to vary greatly over different sets of nuisance parameters when such loci were used.

Outlier SSR spanning coding sequences could be considered candidates for genes under selection. Nonetheless, this should be validated since the risk of false positives is a critical point for all methods that aim to detect a signature of selection (Storz 2005). The simultaneous detection of loci putatively under selection with different tests strengthens the candidate status of a particular locus. In addition, we evaluated the co-localisation of the outlier loci, the genes and quantitative trait loci (QTLs) (Koinange et al. 1996; Blair et al. 2003, 2006a) in order to make inferences about additional hitch-hiking effects.

Genetic population structure

The model-based STRUCTURE (Pritchard et al. 2000) was used to delineate clusters of accessions on the basis of their genotypes at multiple loci, adopting the "admixture model", and the data from neutral and outlier polymorphic SSR loci. Multiple runs of STRUCTURE were performed by setting K (the number of clusters) from 1 to 50. The burn-in time and replication number was set at 50,000 iterations for each run and each run was replicated five times. STRUCTURE generates population structure by maximising the Hardy–Weinberg equilibrium. In order to overcome the fact of the autogamy of common bean, data were transformed to haplotype data assuming a complete homozygosity and were used to generate clusters with the STRUCTURE software. A second assumption of the STRUCTURE is that the loci are unlinked but Pritchard and Wen (2004) indicated that linked data can be used, especially when they are from several linkage groups. Our markers cover all linkage groups in the common bean. The number of clusters (K) was set at the number that maximised the ΔK parameter (Evano et al. 2005), for neutral and outlier loci. Accessions with a membership coefficient of less than 0.8 were identified as putative hybrids. The genetic relationships among entire accessions based on the STRUCTURE for neutral and outlier loci were confirmed by a multiple discriminant analysis using the XLStat programme (http://www.xlstat.com/es/home/). The aim was to predict the likely belonging of an accession to a previously defined qualitative group. A genetic differentiation considering the subsets of accessions based on the cluster STRUCTURE was assessed using the F_{ST} parameter calculated according to the weighted average F statistics over loci (Weir and Cockerman 1984). The level of the gene flow (Nm), which is equivalent to the number of migrants between accessions per generation, was estimated based on

the formula Nm = 0.25 (1 − F_{ST})/F_{ST} (Slatkin and Barton 1989a, b).

Genetic diversity of populations

To avoid making erroneous conclusions from using outlier loci (or from wrongly excluded neutral loci), population parameters were estimated with and without outliers. For polymorphic loci, the mean number of observed and expected alleles, the gene diversity, the heterozygosity, the percentage of loci polymorphic (P), and the polymorphism information content (PIC) were calculated for each microsatellite locus using the PowerMarker 3.25 programme (Liu and Muse 2005). AMOVA within and between STRUCTURE clusters was carried out using the GenAlEx programme (Peakall and Smouse 2006). The genetic relationship among entire populations based on Nei genetic distance (Nei 1978) was analysed by a principal coordinate analysis using the GeneAlEx 6 programme (Peakall and Smouse 2006).

Results

Microsatellite genetic diversity

Out of the 62 SSR loci assayed, 20 were monomorphic overall and these loci were excluded for the subsequent analyses. The overall mean genetic diversity was 0.35 and the average number of alleles per microsatellite locus was 4, ranging from 2 to 11 alleles (BM152). Of the 42 polymorphic loci, 22 (51.16%) (BMd1, BMd8, BMd9, BMd16, BMd17, BMd20, BMd21, BMd33A, BMd3B, BMd37, BM140, BM143, BM152, BM160, BM171, BM183, BM187, BM189, BM197, BM200, BM201 and X04660) showed more than two alleles. The PIC values ranged from 0.01 to 0.82 (BM152) with an average of 0.31 (Table 2). The average gene diversity for genomic SSRs based on 11 microsatellite loci was 0.2812 while that for coding SSRs (20 microsatellite loci) was 0.2753, indicating that the genomic SSRs contributed with more information to the study. The observed heterozygosity was calculated for each individual marker as a measure of marker diversity. The percentage of heterozygotes per marker detected in the common bean landraces ranged from 0% to 38% from marker BM197. The average of observed heterozygosity was 3%, consistent with the predominantly self-pollinated nature of the species.

Table 2 Summary statistics for the 42 polymorphic microsatellite markers analysed in this study

Linkage group	Marker	Number of alleles	Major allele frequency	Gene diversity	Heterozygosity	PIC
1	BM200	8	0.56	0.63	0.013	0.59
2	BMd2	2	0.99	0.01	0.000	0.01
2	BMd7	2	0.91	0.17	0.000	0.15
2	BMd17	3	0.66	0.46	0.000	0.37
2	BMd47	2	0.71	0.41	0.000	0.33
2	BM152	11	0.24	0.84	0.117	0.82
2	BM143[a]	7	0.45	0.65	0.058	0.59
2	BM156	3	0.50	0.52	0.110	0.41
3	BMd1	4	0.49	0.53	0.026	0.42
3	BM197	6	0.35	0.75	0.383	0.71
4	BMd8	3	0.72	0.41	0.000	0.34
4	BMd9	4	0.50	0.57	0.240	0.47
4	BMd15	2	0.98	0.04	0.000	0.04
4	BMd16	3	0.58	0.49	0.000	0.38
4	BMd26	2	0.97	0.06	0.000	0.06
4	BM68	2	1.00	0.01	0.006	0.01
4	BM140	3	0.97	0.05	0.000	0.05
4	BM171	4	0.66	0.51	0.026	0.47
4	X60000	2	0.92	0.14	0.013	0.13
4	X04600	4	0.84	0.28	0.052	0.26
5	BMd20[a]	3	0.96	0.08	0.000	0.07
5	BMd28	2	0.53	0.50	0.000	0.37
5	BMd53	2	0.97	0.05	0.000	0.05
6	BMd12[a]	2	0.82	0.29	0.000	0.25
6	BMd37[a]	3	0.58	0.56	0.000	0.48
6	BM187	7	0.45	0.67	0.000	0.61
7	BMd40	2	0.55	0.49	0.000	0.37
7	BM160	8	0.29	0.78	0.000	0.75
7	BM183	10	0.28	0.83	0.006	0.81
7	BM185	2	0.99	0.01	0.000	0.01
7	BM201	4	0.76	0.39	0.026	0.35
8	BMd25	2	0.99	0.01	0.000	0.01
8	BM189	3	0.97	0.06	0.000	0.06
9	BM188	2	0.72	0.40	0.136	0.32
10	GATS11	2	0.95	0.09	0.019	0.09
11	BMd22	2	0.96	0.07	0.000	0.07
11	BMd33A	3	0.72	0.41	0.000	0.34
11	BMd33B	5	0.82	0.30	0.110	0.27
11	BMd41A	2	0.61	0.48	0.000	0.36
11	BMd41B	2	0.71	0.41	0.019	0.33
Unlinked	BMd21	4	0.48	0.65	0.039	0.58
Unlinked	BMd57	2	0.92	0.14	0.039	0.13
	Mean	4	0.72	0.35	0.034	0.31

[a] Markers with step-wise mutation pattern

Outlier detection

The Beaumont and Nichols (1996) test for neutrality was applied to the 42 polymorphic markers and initially detected 11 outlier loci (mean $F_{ST} = 0.482$) outside the desired confidence interval (0.99% CI). To avoid bias in the differentiation estimate, these highly divergent loci were

removed and after test for neutrality again, no other outliers were found, and a recalculated F_{ST} value (0.440) was used to run the analysis and produce the graphical distribution. In total, five (BMd41A, BMd41B, BM143, BM171 and BM200) and six loci (BMd8, BMd40, BM160, BM183, BM187 and BM197) showed excessively high (exceeding the 0.99 limits) or low F_{ST} values, compared to neutral expectations, which implies that they are good candidates for being subjected to positive and balancing selection, respectively. These loci scattered across seven different linkage groups (B01, B02, B03, B04, B06, B07 and B11, with B07 being the chromosome with the largest numbers of outlier microsatellite loci placed on it), and they could be considered as being affected by local selective pressures.

DetSel pairwise comparisons between populations revealed (31 SSR loci out of 42) BMd1, BMd7, BMd8, BMd9, BMd12, BMd16, BMd17, BMd20, BMd21, BMd22, BMd28, BMd33, BMd37, BMd40, BMd41A, BMd41B, BMd47, BMd53, BMd57, BM143, BM152, BM156, BM160, BM171, BM183, BM187, BMd188, BM197, BM200, BM201, X04660 and X60000 as loci putatively under selection (at $P = 0.99$). The loci BMd26, BM189, GATs11 and BMd212 were discarded as false positives because they were detected in single comparisons. Loci BM152 and BM160 were detected in more than one pairwise comparison, but both were in connection with a particular population (PHA-0419 and PHA-0593), which revealed local selective effects.

After testing both methods, the outlier status of BMd8, BMd40, BMd41, BM143, BM160, BM171, BM183, BM187, BM197 and BM200 was confirmed. It is not surprising that both tests identified different outlier loci from the same data set, because they use different summary statistics: LOSITAN uses F_{ST} (the standardised variance of allele frequencies) and DetSel uses F (an estimate of identity by descent—the probability that two alleles sampled in a population are identical owing the shared ancestry). BM160 was revealed as a potential locus under selection in the PHA-0593 population and under both LOSITAN and DetSel scenarios. BM183 was affected by selection in the PHA-1068 population and under the LOSITAN scenario only. The outlier SSRs detected in this study are also co-localised with important QTLs and genes.

Population structure

Genotype data for 31 neutral SSR markers were used to determine the population structure among the various common bean accessions. After applying the STRUCTURE software (Fig. 1), the identification of the genetic origin for each accession was accomplished as described in 'Materials and methods' section for $K = 3$. For $K = 3$, 91.1% accessions were strongly assigned to one cluster or another and the proportions assigned to each group were asymmetric, which strongly indicated that there existed a real population structure. Individuals were assigned to specific populations or multiple populations (if their genotypes indicated admixture) based on their membership coefficients. When the microsatellite data set was trimmed to include all loci markers, not similar results were obtained (data not shown). $K = 4$ and 6 were also suggested as possible groups of clustering after calculating the ΔK value, and the frequency of accessions, which could not be clearly assigned to any one of these clusters was similar (from 11.9 to 13.7%) to the one in $K = 3$.

The summary statistics of the neutral microsatellite diversity in the clusters identified by STRUCTURE for $K = 3$, 4 and 6 is shown in Table 3. The genetic diversity in the cluster of North-Eastern accessions at $K = 3$ was slightly higher than in the other Mesoamerican population clusters (0.27 and 0.22, respectively), although that cluster had the smallest sample size in this study. However, a reduction in genetic diversity of about 10%, whether measured by genetic diversity or PIC values, was observed in different STRUCTURE clusters. For $K = 3$, the cluster of Mesoamerican check cultivars (included in cluster 2) showed a mixed membership among the three clusters. The Andean check cultivar is only presented with North Eastern accessions (included in cluster 2) at $K = 3$ and 4 while that at $K = 6$ is separated from the entire collection of accessions, giving rise to the cluster C6.

A multiple discriminant analysis based on the clustering achieved by the STRUCTURE programme, confirmed the population structure suggested by the application of this software (Fig. 2). The principal component analysis based on Nei distance allowed us to separate similar clusters of accessions as per STRUCTURE (Fig. 3). This split is in good accordance with planting ecotypes and geographical sowing zones of the cultivated common bean in the south of Europe. The geographical characteristics of each cluster evidenced in Fig. 2 were as follows:

Cluster 1 contained five landraces, of which four (80.39%) were accessions from the *Feijao manteiga* genotype cultivated in the northern provinces of Portugal (41°N, 7°W), one landrace (19.6%) from the *Plancheta* type cultivated in Leon and the Mesoamerican Almonga cultivar. In addition, summer common bean/maize cropping system is the main cultivation system in those high level areas (~600 masl). This study has also allowed us to quantify the population admixture for each population. The proportion of non-hybrid accessions was 67.3% at the 0.8 cut-off value.

Cluster 2 comprised accessions from six landraces, of which one (12.2%) was from the *Faba pancha* genotype and is cultivated at low high level planting (20 masl) in the north of Spain; two landraces (37.5%) were from the

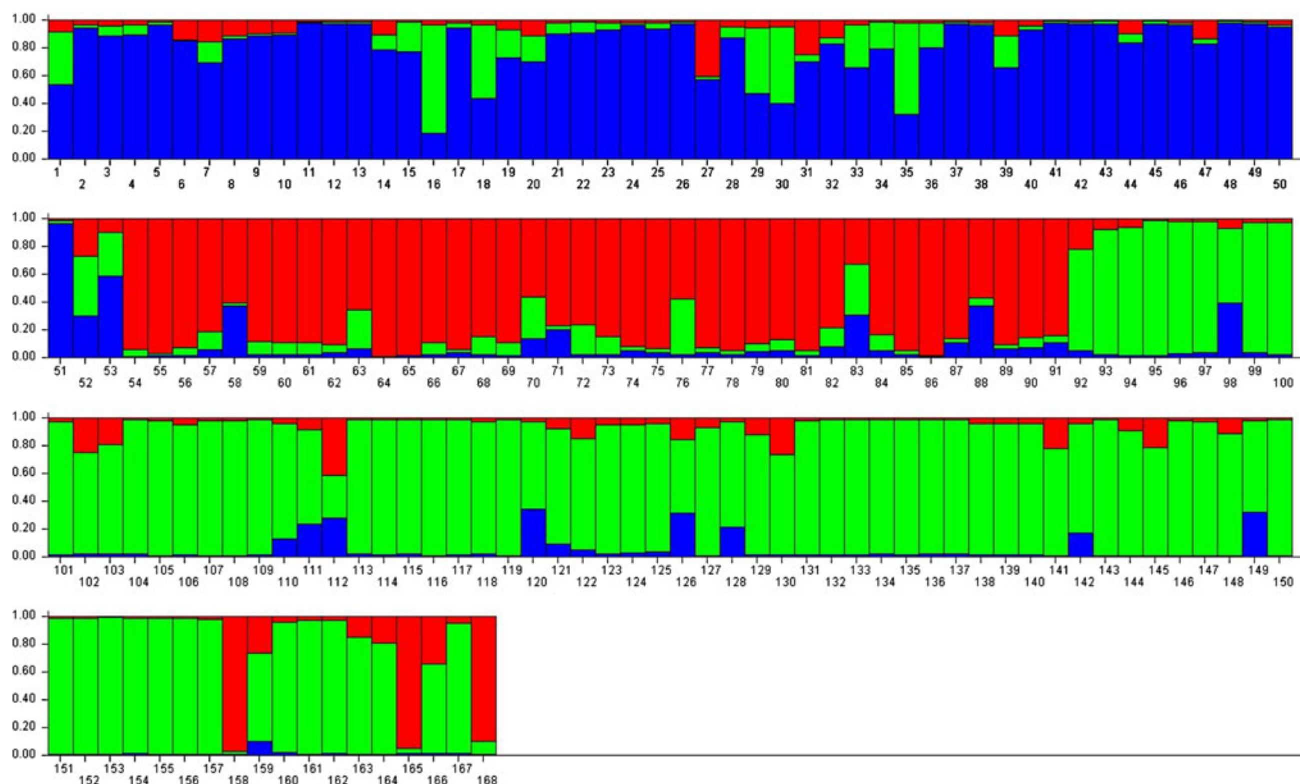

Fig. 1 Hierarchical organisation of genetic relatedness of 154 accessions from 22 common bean landraces based on 31 neutral microsatellites. Each individual bar represents an accession which is partitioned into $K = 3$ coloured segments. Each *colour* represents one cluster, and the length of the coloured segment shows the accession's estimated proportion of membership in that cluster as calculated by the STRUCTURE at that value of K. Numbers 1–41 and 92–104, Feijao manteiga; 42–51 and 149–160, Planchada and Plancheta; 52, Alert; 53, Almonga; 54, Beryl; 55, ICA Pijao; 56, Matterhorn; 57, UI465; 58, USGN-5; 59–61, USWA12; 62, USWA13; 63, Weihing; 64, Calima; 65, MDRK; 66-71 and 77–85, 109–112 and 120–125, Garabanzo; 86–91, Ganxet; 105–108, 113–119 and 157–168, Faba do caldo; 126–141, Chichos; and 142–148 = Habichuela accessions

Garabanzo genotype, which are cultivated in the north of Spain (42°N, 9°W), and two landraces (15%) displayed a *Ganxet* genotype, which are usually cultivated in the northeastern region (41°N, 2°W) of Spain. It also included the Great Northern control cultivars Alert, Beryl, Matterhorn, UI465, USGN-5, USWA12, USWA13, Weihing and ICA Pijao, and the Andean cultivars Calima and Michigan Dark Red Kidney. The proportion of non-hybrid accessions was 78.1% at the 0.8 cut-off value.

Cluster 3 included accessions from 13 landraces, of which 3 (29.9%) belong to the *Faba do caldo* type cultivated in the north of Spain (43°N, 7°W) at lowland fields (~20–118 masl), three landraces (16.8%) were from the type *Feijao manteiga* cultivated in the north of Portugal (41°N, 7°W), two landraces (10.4%) were from the *Planchada* and *Plancheta* genotypes cultivated in highland fields (~800 masl) of the Bañeza region (43°N, 5°W) in the north of Spain, two landraces (12.9%) were from the *Garabanzo* type cultivated in the growing area (42°N, 8°W) in the north of Spain, one landrace (20.8%) was a *Chichos* type, and finally, one landrace (9.1%) belongs to the *Habichuela* type. The proportion of non-hybrid accessions was 78.7% at the 0.8 cut-off value.

On average, the F_{ST} value among clusters at $K = 3$ was 0.090 with F_{ST} for each locus ranging from 0.006 (locus BMd22) to 0.254 (locus BMd21). AMOVA analysis showed an increase in the genetic differentiation for Mesoamerican common bean clusters from $K = 3$ ($F_{ST} = 0.090$) to $K = 4$ and 6 ($F_{ST} = 0.163$ and 0.178, respectively; Table 4). F-statistic estimates over all the neutral loci showed an average of inbreeding in the total of clusters and within the population coefficient of 0.896 at $K = 3$, and 19 loci showed complete fixation. Pairwise comparison on the basis of the values of F_{ST} can be interpreted as standardised population distances between two populations (Li and Nelson 2001; Chen and Nelson 2005). Presumably divergent accessions from the northeast of Spain cluster had the highest F_{ST} value with all the clusters of accessions from Portugal and Spain (average pairwise $F_{ST} = 0.167$; Table 5). For instance, Nm (a combined estimate of the effective population size and migration rate) was estimated to be 2.514 when the three clusters are considered in the

Table 3 Summary statistics of the neutral microsatellite diversity in the common bean population clusters identified by the STRUCTURE software for $K = 3$, 4 and 6

Clusters (C)	Sample size	Average no. alleles	Average no. effective alleles	Polymorphic loci percentage	I	Gene diversity	Heterozygosity	PIC
$K = 3$								
North of Portugal accessions (C1)	51	2.424	1.459	78.79	0.403	0.223	0.031	0.2084
Northeast-west of Spain accessions (C2)	29	2.364	1.566	87.88	0.455	0.270	0.030	0.2068
Northwest of Spain accessions (C3)	74	2.061	1.458	72.73	0.379	0.225	0.017	0.1976
Total	154	2.283	1.494	79.80	0.412	0.240	0.026	0.2346
$K = 4$								
North of Portugal accessions (C1)	47	2.424	1.476	78.79	0.408	0.226	0.031	0.2003
North West of Spain accessions (C2)	42	1.848	1.342	60.61	0.281	0.161	0.011	0.2084
Northwest of Spain accessions (C3)	55	2.303	1.514	81.82	0.421	0.251	0.026	0.1651
Northeast of Spain accessions (C4)	10	1.788	1.437	57.58	0.365	0.229	0.033	0.1893
Total	154	2.091	1.442	69.70	0.369	0.217	0.025	0.2346
$K = 6$								
Northeast of Spain accessions (C2)	11	1.758	1.390	57.58	0.342	0.213	0.035	0.1656
North of Portugal accessions (C3)	48	2.394	1.466	78.79	0.404	0.224	0.030	0.2066
Northwest of Spain accessions (C4)	34	2.182	1.468	78.79	0.387	0.229	0.030	0.1884
Northwest of Spain accessions (C5)	24	1.788	1.273	51.52	0.258	0.147	0.015	0.1299
Northwest of Spain accessions (C6)	37	1.970	1.428	67.27	0.351	0.211	0.014	0.1800
Total	154	2.018	1.405	67.27	0.348	0.205	0.025	0.2346

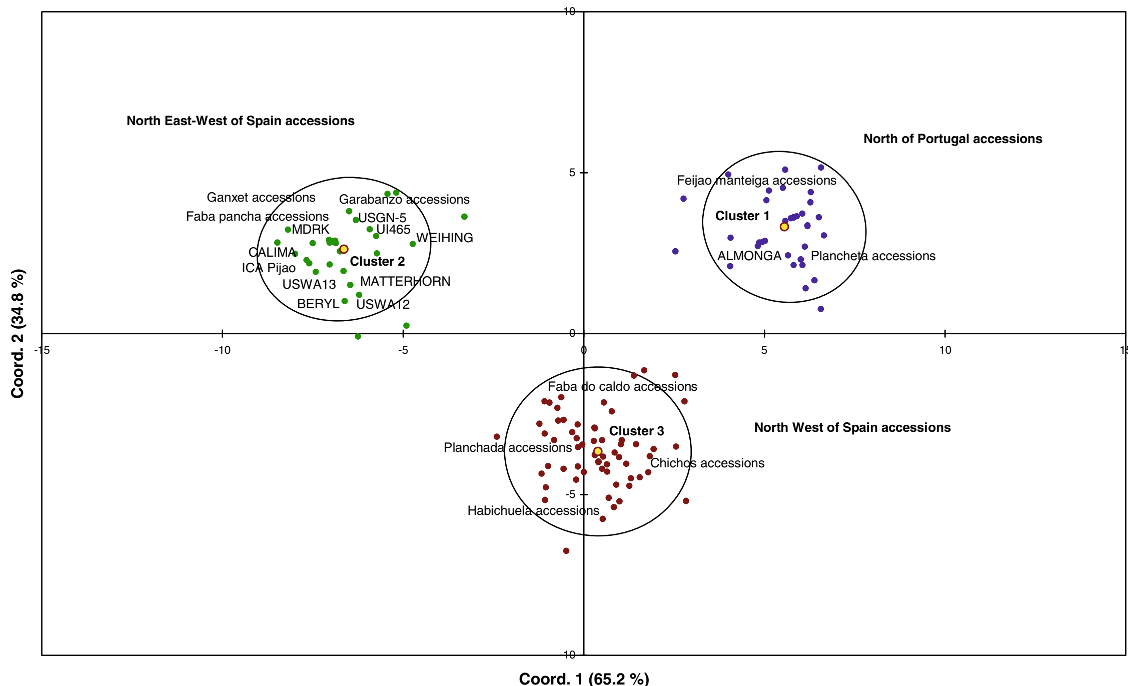

Fig. 2 Multiple discriminant analysis of the neutral microsatellite diversity based on the presence and absence of alleles. *Colours* represent populations identified at $K = 3$ in Fig. 1

analysis (e.g. if the average Ne is 100, the migration rate per generation is 0.0486) (Crow and Aoki 1984).

A similar population structure was uncovered with the principal component analysis (Fig. 3), in particular the subdivision between the presumably divergent accessions as *Ganxet* type from the other groups of ancestral accessions as *Faba pancha* and *Feijao manteiga* types, which were well separated on the principal coordinate 1 (45.0%).

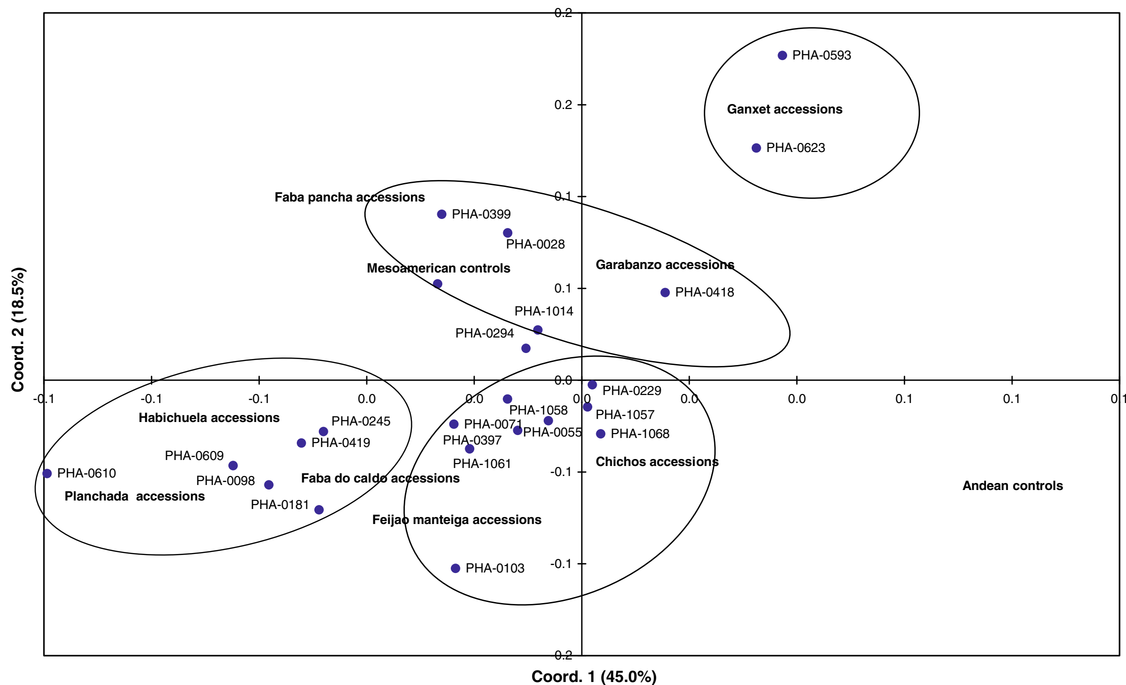

Fig. 3 Principal component analysis of the neutral microsatellite diversity of Mesoamerican genotypes of the cultivated common bean

Table 4 Results of the AMOVA in common bean clusters of populations identified by the STRUCTURE software for $K = 3$, 4 and 6 based in the neutral microsatellite diversity

Cluster and source of variation	SS	Variance	Total (%)
Genetic origin clusters $K = 3$ ($F_{ST} = 0.090$** Nm = 2.514 $F_{IS} = 0.896$ $F_{IT} = 0.896$)			
Among clusters	305.114	3.139	17
Within clusters	1992.428	15.094	83
Total	2297.543	18.233	
Genetic origin clusters $K = 4$ ($F_{ST} = 0.163$** Nm = 1.281 $F_{IS} = 0.880$ $F_{IT} = 0.895$)			
Among clusters	444.100	4.244	23
Within clusters	1852.903	14.144	77
Total	2297.003	18.389	
Genetic origin clusters $K = 6$ ($F_{ST} = 0.178$** Nm = 1.155 $F_{IS} = 0.871$ $F_{IT} = 0.893$)			
Among clusters	538.549	4.634	26
Within clusters	1758.468	13.527	74
Total	2297.07	18.161	

SS sum of squares: % total refers to the percentage of total variance contributed by each component

** $P \leq 0.01$

Furthermore, a small group of accessions from the north of Spain was positioned among the group of divergent accessions and the other two large groups of accessions according to the principal coordinate 2 (18.5%), but skewed towards the Mesoamerican controls.

Phenotypic diversity

To consider the structure of the molecular diversity that exists among the accessions studied according to the loci affected for selection, a model-based analysis was performed using the STRUCTURE software and the data from 10 outlier SSR loci. After setting the number of clusters according to the statistics of Evano et al. (2005) to three, the software outputs the coefficient of estimated ancestry per each accession in each cluster. The plot of ancestry estimates shown in Fig. 4a in parallel with the multiple discriminant graph (Fig. 4b) based on the morpho-physiological data, is broken into three segments, with lengths proportional to the genotype's estimated ancestry fraction from each of the three clusters. Model-based groups were mostly consistent with the phenotypic classification. The clusters that, based on the molecular and phenotypic diversity, contained the accessions were plotted in a plane defined by the first two axes which accounted for 100% of the total variation (Fig. 4b). This representation evidenced three clusters with particular phenotypic characteristics and highlighted a relation between the origin due to selection and the phenotypic diversity and thus confirmed the clustering generated by the STRUCTURE software. Seeds of landraces representative of each cluster are documented in Fig. 4d (I–III).

The first SSR-based cluster contained the accessions that have a hooked seed shape and belong to the *Ganxet* (*Ganxet* means a "little hook" in Spanish) bean market

Table 5 F_{ST} (below the diagonal) and Nm (above the diagonal) values between all pairwise combinations of common bean clusters identified by the STRUCTURE for $K = 3$, 4 and 6

$K = 3$	C1	C2	C3		
North of Portugal accessions (C1)	0.000	3.072	3.495		
Northeast-West of Spain accessions (C2)	0.075	0.000	3.785		
Northwest of Spain accessions (C2)	0.067	0.062	0.000		
$K = 4$	C1	C2	C3	C4	
North of Portugal accessions (C1)	0.000	2.567	3.144	1.406	
Northwest of Spain accessions (C2)	0.089	0.000	2.600	1.225	
Northwest of Spain accessions (C3)	0.074	0.088	0.000	1.622	
Northeast of Spain accessions (C4)	0.151	0.169	0.134	0.000	
$K = 6$	C1	C2	C3	C4	C5
Northeast of Spain accessions (C1)	0.000	1.669	1.785	1.669	1.273
North of Portugal accessions (C2)	0.130	0.000	2.563	2.383	2.383
Northwest of Spain accessions (C3)	0.123	0.089	0.000	1.758	2.380
Northwest of Spain accessions (C4)	0.130	0.095	0.125	0.000	2.063
Northwest of Spain accessions (C5)	0.164	0.095	0.095	0.108	0.000

All pairwise F_{ST} values are significant at $P \leq 0.001$

class cultivated in the northeast of Spain, and the accessions that have a flattened rhombohedral shape from the *Planchada* market class cultivated in the northwest of Spain, both greatly appreciated for their organoleptical qualities. These accessions flower and mature later, have large chordate leaflets, long fifth internodes, fruit commencing at higher nodes, a lower yield per plant, and a higher incidence of resistance (scores <4) to the Mesoamerican (17, 73, 448 and 1545) and Andean (7, 39 and 102) races of anthracnose, and reaction to the BCMNV. These accessions presented the $Skdh^{103}$ allele and the S phaseolin pattern, and the phenotype resembles the typically Andean germplasm.

The second SSR-based cluster corresponding to accessions exhibited a shorter fifth internode, flowering commencing from and concentrated in the basal nodes, high pods per plant, early flowering, large oval leaflets, small seeds, and a higher incidence of susceptibility (scores >4) to the Mesoamerican (7, 448 and 1,545) races of anthracnose and resistance (scores <4) to the Andean (7, 39 and 102) races of anthracnose and reaction to the BCMV. This group is characterised by the presence of the $Skdh^{100}$ allele and the B (20%) phaseolin pattern.

Finally, the co-ancestry analysis formed a cluster that included the large seeded accessions belonging to the *Faba pancha* and *Feijao manteiga* bean market class cultivated in the northwest of Spain and Portugal, respectively. Therefore, this was also paralleled to the analysis based on the phenotypic variability. Accessions belonging to this group exhibited earlier flowering commencing from and concentrating on the basal nodes, long fifth internode lengths, small to medium oval leaflets, a higher yield per plant, large seeds, and a higher incidence of susceptibility (scores >4) to the Mesoamerican (7, 448 and 1,545) races of anthracnose, and resistance (scores <4) to the Andean (39 and 102) races of anthracnose, halo blight and reaction to the BCMNV. This group is also characterised by the presence of the $Skdh^{103}$ allele and the B (20%) phaseolin pattern. The genotypes that belong to this group fit the general description of the race Durango of the Mesoamerican gene pool.

The analysis of molecular diversity affected by selection reveals that, within the SSR-based groups, there is uniformity at the phenotypic level. The clusters of Mesoamerican landraces with similar outlier SSR pattern exhibited similar morphology adaptation and phenology. In addition, most individuals from the same landrace tended to group together, except for landraces PHA-0294, PHA-0399, PHA-0419, PHA-1057, and PHA-1058 with individuals grouped in clusters 1 and 3, landraces PHA-0397, PHA-0418, PHA-1014, and PHA-1061 with individuals in clusters 1 and 2, and landrace PHA-0229 with individuals grouped in clusters 3 and 2.

Discussion

Phenotypic and non-neutral molecular diversity

The detection of outlier loci consisted in screening many loci scattered in the genome to bring out the few that diverged from empirical or simulated neutral expectations (Luikart et al. 2003). Compared with the rest of the genome, the outlier loci have an atypical behaviour, which can range from an excess or a deficit of rare alleles in a given landrace to aberrant patterns of genetic variability within or among landraces (Luikart et al. 2003; Schlötterer 2003; Nielsen 2005). Because revealing outlier loci in genome scans currently depends on statistical tests, one of the main concerns is to highlight truly significant loci while avoiding the detection of false positives as much as possible.

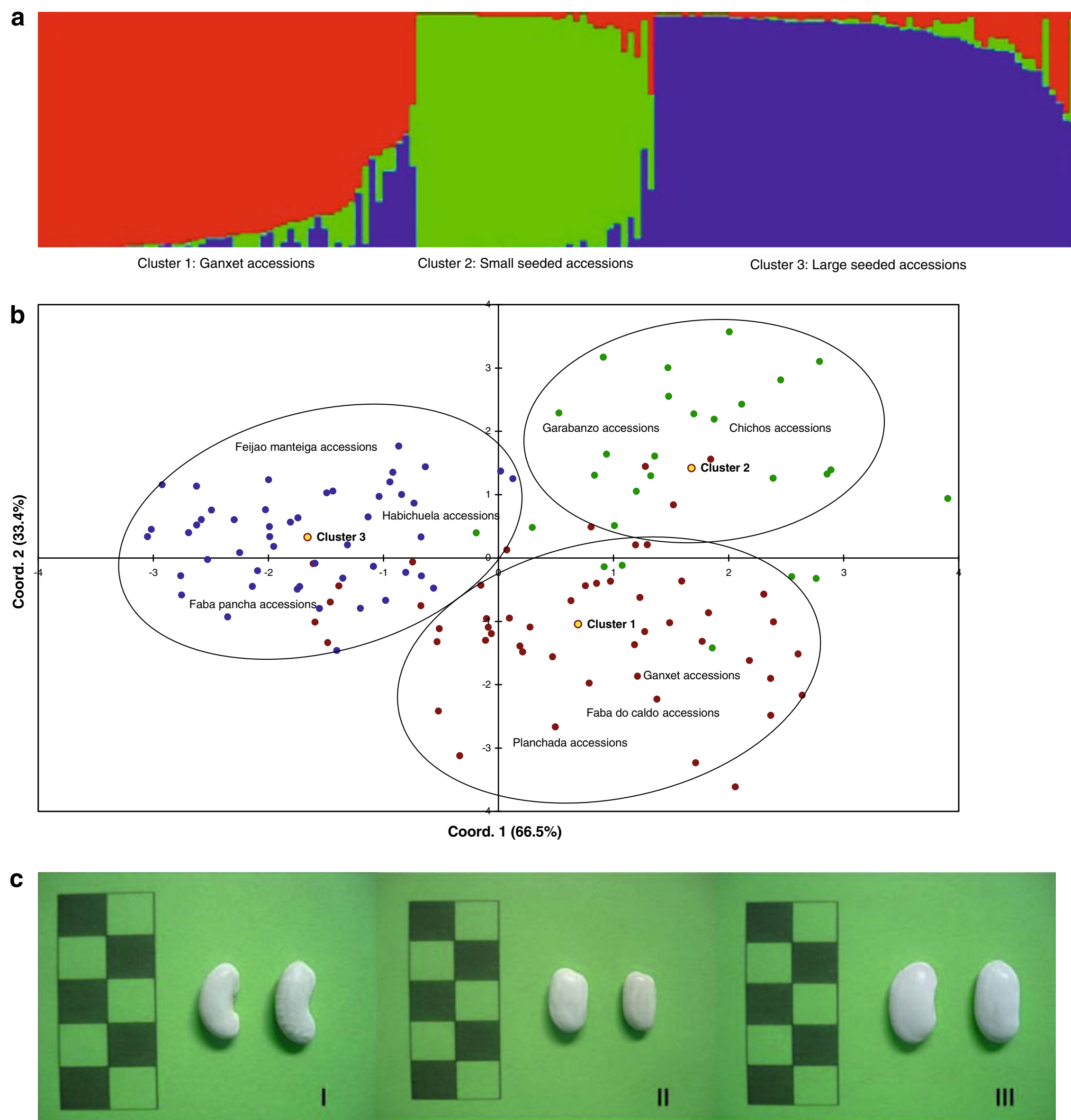

Fig. 4 Phenotypic and molecular analysis based on 10 outlier microsatellites of accessions from 22 common bean landraces belonging to the Mesoamerican gene pool. **a** Estimated population structure; each individual is represented by a *horizontal bar*, which is partitioned into three coloured segments that represent the individual estimated levels of the three clusters. **b** Multiple discriminant analysis based on phenotypic traits (*coloured symbols* highlight the major clusters). **c** Images of seeds collected from landraces that are representative of the seed type variations (I, cluster 1-Ganxet accessions; II, cluster 2-small seeded accessions; III, cluster 3-Large seeded accessions). *Bar* 5 cm

The methods used in this work agree on considering loci BMd8, BMd40, BMd41, BM143, BM160, BM171, BM183, BM187, BM197 and BM200 as good candidates for selection. Thus, if these outlier loci were important in the adaptation of the common bean to the new environmental conditions in Europe, their continuous manipulation might contribute to future genetic gain with which breeders could work. Data relating to the phenotypic characterisation considered along with the results of the outlier molecular structure appear to support this hypothesis.

The outlier SSRs detected in this study are also co-localised with important QTLs and genes mapped by Yu et al. 2000, Gaitán-Sólis et al. 2002, Blair et al. 2003, 2006a and Pañeda et al. 2008. Consequently, the non-neutral behaviour of these loci may alternatively indicate selective effects at other loci via hitchhiking. Farmer-intentional (for large seed size, high yield or for different flowering dates) or non-intentional (for disease resistance) selective pressures may underline these effects. It is also observed that loci with a significantly higher differentiation tend to gather in the

same chromosomal locations (Bonin et al. 2006) as loci BMd40, BM160 and BM183 in the B07 linkage group. Translating this information into tools for marker-assisted breeding would then be straightforward.

Overall, morphological and molecular data showed that the 164 bp allele of the locus BM183, and the 184 and 228 bp alleles of the locus BM160, were specific alleles for cluster 1, the group of accessions with the highest resistance to the Mesoamerican races (17, 448 and 1,545) of anthracnose and the highest number of nodes to flowering. Both loci are located on the chromosome B07 at a location that is flanking plant height *ph7.1* and the *Co-5-Co-6* anthracnose resistance gene sequences and thus they might be in disequilibrium with those genes involved in the plant defence (Blair et al. 2006b; Pañeda et al. 2008).

Accessions belonging to clusters 1 and 2 had the largest leaf size and presented the private 168 and 400 bp alleles, respectively, of the locus BM187 located in the genomic region of a QTL for leaf size on the chromosome B06. A high frequency of the specific 263, 239 and 253 bp alleles of the locus BM200 was observed in each group of accessions, which had resistance to the Mesoamerican race 73 of anthracnose and different earliness. This locus is situated on the chromosome B01 in the genomic region of the *Fin* gene for growth habit and the *Co-2* anthracnose resistance gene (Pañeda et al. 2008).

Specific 92 and 148 bp alleles of the locus BM143 were observed in clusters 1 and 3, including the accessions with the highest resistance to BCMNV, and also differed for earliness and seed size. This locus is located on the chromosome B02 at a location that is close to the seed weight *sw2.1*, days to flowering *df2.1* and the BCMV resistance gene *I* (Blair et al. 2006b).

The private 135 bp allele of the BM171 locus was found in a high frequency in cluster 3, accessions that had the highest yield per plant and internode length, and resistance to the Mesoamerican race 73 of anthracnose. This locus is located in the genomic region of several QTLs for yield *yld4.2* and *yld4.3* and internode length *Int3*, *Int2* and *Int4*, and gene sequences conferring resistance to race 73 of anthracnose included in the *Co-3*/*Co-9* cluster on the chromosome B04 (Checa and Blair 2008; Pañeda et al. 2008).

These studies can identify these outlier loci but they cannot demonstrate whether they are positioned or not. Moreover, they were unable to establish a definite marker-morphological trait relationship. Conclusive evidence of selection can only come from functional studies where fitness-related variation among alternative genotypes is measured directly, as suggested by Storz (2005). Thus, these supposedly selected loci will have to be considered with caution until corroborating evidence is acquired proving that the selection is really responsible for their outlier nature. From an examination of the seed weight, it was evident that, the Mesoamerican accessions studied were selected for large seed size. Therefore, it is feasible that different farmer-selection pressures applied for generations, on the same probably restrictedly introduced gene pools, have had a significant impact on the divergence loci among accessions. Earliness and seed size still played an important role in the obtained classification. These traits are highly heritable and can be easily modified by selection, so that bean accessions with the same genetic background may have different earliness or seed sizes. As a consequence, non-neutral microsatellite analysis does not seem fully appropriate for the classification of bean landraces according to their genetic origin. It has been demonstrated that major phenotypic differences can often occur with only minor genotypic changes (Rick and Holle 1990). Nevertheless, our findings are consistent with the hypothesis that different selective forces, selection by farmers, adaptation to heterogeneous environments and interactions with other species such as parasites are involved in maintaining the diversity of local common bean landraces studied.

Genetic population structure

The genetic diversity detectable in crops by molecular markers such as SSRs is generally dependent on the mating system, the domestication history, and the magnitude of the collection being analysed. In self-pollinating species, as the polymorphism is typically maintained among accessions, the estimated genetic diversity results in a function of the sample size. When wide collections are addressed (e.g. more than 100 entries), numbers of alleles per polymorphic locus above 10 have commonly been reported in studies of self-pollinating crops, such as wheat (Huang et al. 2002), rice (Xu et al. 2005), and soybean (Wang et al. 2006).

Since the first studies of the common bean using molecular markers, it was clear that the level of genetic diversity in the cultivated gene pools was significantly lower than in other self-pollinating species. The analysis of 60 genotypes representing a diverse array of American Mesoamerican germplasm with 52 SSR loci, Díaz and Blair (2006) found an average of 5.1 alleles per locus for the average of Mesoamerican races, although this varied from 2.8 for the race Durango to 3.2 for the race Mesoamerica. In general, the screening of large sample sets of wild and domesticated common bean cultivars yields a high number of alleles per loci (Kwak and Gepts 2009), something which is not expected when working with low sample sets. However, we demonstrated in this study that a relatively small number of diverse landraces from a reduced geographical area can encompass similar values of genetic diversity as the one shown in large collections. This is supported by the fact that we scored a number of alleles per polymorphic locus of 4, taking into account the 154 accessions of 22 cultivated

common bean landraces. In this respect, the total gene diversity across all accessions in our study was similar to that reported in previous studies for the races Durango and Mesoamerica (Díaz and Blair 2006). In spite of achieving relevant results, these comparisons should be managed with caution, because all these values can be altered depending on both the sample size and the SSRs selected for the analysis. According to this, only general conclusions can be disclosed from these comparisons. Whole analysed samples in our study belong not only to the same race, but also to the same commercial type and therefore, for such a homogeneous group of accessions, non-significant values in gene diversity are expected. Unlike this, they appeared to be a set of samples with both high total numbers of alleles per locus and high gene diversity values. So, in light of these results and under the above considerations, it can be suggested that our study confirmed the broad genetic base of the Mesoamerican accessions of the common bean from southern Europe. This may well be expected because Spain and Portugal are regarded as a centre of secondary diversification for the common bean (Santalla et al. 2002).

Many reports have been published that describe differences for F_{ST} at different loci. The results have revealed a high ($F_{ST} = 0.178$) differentiation that is cogent with the absence of a gene flow. In addition, the high value of genetic differentiation among accessions suggests a high autogamy or genetic drift (Zizumbo-Villarreal et al. 2005). As highlighted by the graphic representation, this difference is mostly attributable to the contribution of the North East landraces (*Ganxet* accessions), which also holds the highest overall values of pairwise F_{ST} values and the lowest gene flow values. On the other hand, if these divergent accessions are removed from the analysis, the F_{ST} drops and the Nm increases indicate that they are the most divergent accessions, as was previously shown using the allelic frequencies and placed in the genetic distances. It is not certain whether this high level of heterogeneity was fostered by the action of qualitative distinct selective forces in different areas, as superficial observations of the habitats have revealed no obvious ecological differences, and different tests have indicated that the loci employed can be considered neutral. Furthermore, although the geographic isolation between the east and the west of the Iberian Peninsula is a partial barrier to the gene flow, the detected levels of differentiation are not only due to this factor, as genetic heterogeneity within groups is also substantial. In the face of this evidence, our genetic results are indicative that habitat fragmentation has led to a critical situation where the influence of the genetic drift associated with a low population size and insufficient inter-population interchange have affected the genetic structure of European populations belonging to the Mesoamerican gene pool. This is an important fact to bear in mind as breeders and geneticists search for unique genes or gene combinations among genotypes.

The region with the highest genetic diversity is generally considered as the centre of the origin or secondary domestication of a species (Vavilov 1951). Out of all the clusters, those from the north of Portugal and northwest of Spain (Galicia and Asturias) had the highest allelic richness and the largest number of unique (i.e. cluster-specific) alleles. Galicia, Northern Portugal, Asturias, Western León, and Zamora have formed a single area since the Neolithic and Chalcolithic (also called Copper Age) Ages, around 4,500–1,500 BCE, and it is estimated that fewer than 30,000 people have settled there, mainly in the urbanised zones of Portugal (Braga and Porto), Lugo and León. Furthermore, one of the main Roman routes starts in Porto (Portugal) and enters Spain via Tui. Our results hypothesised that this geographical area may be considered a candidate secondary centre of the common bean diversity (Santalla et al. 2002). The high genetic diversity within these accessions may be the result of several factors such as diverse environments and multiple types of usages. This diversity is reflected in different terms in the Spanish-speaking of this area: the common bean may be called garabanzo, faba and feixon (from the Latin word "Phaseolus"), and feixo, which is a common family name in the northwest of Spain.

Conclusions

The STRUCTURE and multivariate analyses of the SSR data are approaches with different assumptions which have led to a deeper understanding of the relationships of common bean landraces and their origin. Moreover, the results complement one another and provide a robust analysis. STRUCTURE utilises a Bayesian clustering approach to assign individuals probabilistically to populations based on their genotypes in attempts to find populations structure. The results showed that the structure of the genetic diversity in the Mesoamerican common bean accessions studied depends on the nature of the SSR polymorphisms used. A contrast is observed in this work between the selectively neutral or non-neutral nature of microsatellites. Non-neutral microsatellites are subject to both natural and artificial selection since environmental conditions and the farmer's selection criteria led to divergence among genotypes. This work supports the superiority of neutral molecular data for defining groups of landraces with similar origin. It has been also demonstrated that common bean landraces that are still growing in Spain and Portugal maintain a high level of genetic diversity. Therefore, the best way to classify and describe genetic resources such as traditional bean landraces appears to be a two-step process: firstly, a classification based on neutral molecular data, and secondly, a

morphological description of each group based on an outlier molecular approach.

Acknowledgments Research was supported by RF2006-00001, AGL2008-02091/AGR and INCITE07PXI403088ES projects from the Spanish Government and Xunta de Galicia, respectively. The authors are grateful to the Diputación Provincial de Pontevedra for farm facilities and María Lores for technical assistance. M. De La Fuente is grateful to the Xunta de Galicia for awarding her a fellowship grant, which allowed her to carry out this study. Special thanks to J.D. Kelly for anthracnose, S.P. Singh and P. Miklas for BCMV and BCMNV, and D. Fourier and J. Murillo for bacterial inocula.

References

Afanador L, Hadley S, Kelly JD (1993) Adoption of a mini-prep DNA extraction method for RAPD marker analysis in common bean (*Phaseolus vulgaris* L.). Bean Improv Coop 36:10–11

Antao T, Lopes A, Lopes RJ, Beja-Pereira A (2008) LOSITAN: A workbench to detect molecular adaptation based on a *Fst*-outlier method. BMC Bioinformatics 9:323

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

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. Proc R Soc B 263:1619–1626

Beaver JS, Kelly JD (1994) Comparison of selection methods for dry bean landraces derived from crosses between gene pools. Crop Sci 25:923–926

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

Beebe S, Toro O, González AV, Chacón MI, Debouck D (1997) Wildweed-crop complex 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

Blair MW, Pedraza F, Buendia H, Gaitan E, Beebe S, Gepts P, Tohme J (2003) Development of a genome wide anchored microsatellite for common bean (*Phaseolus vulgaris* L). Theor Appl Genet 107:1362–1374

Blair MW, Giraldo MC, Buendia HF, Tovar E, Duque MC, Beebe S (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 landrace derived from a cultivated Andean wild common bean (*Phaseolus vulgaris* L) cross. Theor Appl Genet 112:1149–1163

Blair MW, Fregene MA, Beebe SE, Ceballos H (2007) Marker-Assisted Selection in Common Beans and Cassava. Chap. 7. In: Guimaraes E (ed) Marker-assisted selection (MAS) in crops, livestock, forestry and fish: Current status and the way forward. FAO, Rome, Italy

Bonin A, Taberlet P, Claude M, Pompanon F (2006) Explorative genome scans to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). Mol Biol Evol 23(4):773–783

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

Buso GSC, Amaral ZPS, Brondani RPV, Ferreira ME (2006) Microsatellite markers for the common bean - *Phaseolus vulgaris*. Mol Ecol Notes 6:252–254

Caixeta ET, Borem A, Alzate-Marin AL, Fagundes S (2005) Allelic relationships for genes that confer resistance to angular leaf sport in common bean. Euphytica 145:237–245

Campbell D, Bernatchez L (2004) Generic Scan Using AFLP Markers as a Means to Assess the Role of Directional Selection in the Divergence of Sympatric Whitefish Ecotypes. Mol Biol Evol 21(5):945–956

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

Checa C, Blair MW (2008) Mapping QTL for climbing ability and component traits in common bean (*Phaseolus vulgaris* L.). Molecular Breed 22:201–215

Chen Y, Nelson RL (2005) Relationship between origin and genetic diversity in chinese soybean germplasm. Crop Sci 45:1645–1651

Crow JF, Aoki K (1984) Group selection for a polygenic behavioral trait: estimating the degree of landrace subdivision. Proc Natl Acad Sci USA 81:6073–6077

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

Díaz LM, Blair MW (2006) Race structure within the Mesoamerican gene pool of common bean (*Phaseolus vulgaris* L.) as determined by microsatellite markers. Theor Appl Genet 114:143–154

Evano G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

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

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

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

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 (1988) Phaseolin as an evolutionary marker. In: Resources of *Phaseolus* beans. Kluwer, Dordtrecht, pp 215–241

Gepts P (2004) Domestication as a long-term selection experiment. Plant Breed Rev 24(part 2):1–44

Gepts P, Bliss FA (1985) F1 hybrid weakness in the common bean: differential geographic origin suggests two gene pools in cultivated bean germplasm. J Hered 76:447–450

Gepts P, Bliss FA (1988) Dissemination pathways of common bean (*Phaseolus vulgaris*, Fabaceae) deduced from phaseolin electrophoretic variability. II. Europe and Africa. Econ Bot 42:86–104

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

Guerra-Sanz JM (2004) Short communication - New SSR markers of *Phaseolus vulgaris* from sequence databases. Plant Breed 123:87–89

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

Huang XQ, Börner A, Röder MS, Ganal MW (2002) Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. Theor Appl Genet 105:699–707

Innes NL, Conway J, Taylor JD (1984) Resistance to halo blight in the Cambridge accessions V4604 and V4058 of *Phaseolus* beans. Ann Appl Biol 104:307–314

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

Jarvis DI, Hodgkin T (1999) Wild relatives and crop cultivars: detecting natural introgression and farmer selection of new genetic combinations in agro-ecosystem. Mol Ecol 8:S159–S173

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

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

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

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

Kwak M, Gepts P (2009) Structure of genetic diversity in the two major gene pools of common bean (*Phaseolus vulgaris* L., Fabaceae). Theor Appl Genet 118:979–992

Li Z, Nelson RL (2001) Genetic diversity among soybean ac cessions from three countries measured by RAPDs. Crop Sci 41:1337–1347

Liu K, Muse S (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics Appl Note 21(9):2128–2129

Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA (2006) Phylogeography of Asian wild rice, Oryza rufipogon, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. Proc Natl Acad Sci USA 103:9578–9583

Luikart G, England PR, Tallmon D (2003) The power and promise of population genomics: From genotyping to genome typing. Nat Rev Genet 4(12):981–994

Masi G, Logozzo P, Donini P, Spagnoletti Z (2009) Analysis of genetic structure in widely distributed common bean landraces with different plant growth habits using SSR and AFLP markers. Crop Sci 49:187–199

McClean PE, Lee RK, Miklas PN (2004) Sequence diversity analysis of dihydrofl avonol 4-reductase intron1 in common bean. Genome 47:266–280

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

Mitchell SE, Kresovich S, Jester CA, Hernandez CJ, Szewc-McFadden AK (1997) Application of multiplex PCR and fluorescence-based, semi-automated allele sizing technology for genotyping plant genetic resources. Crop Sci 37:617–624

Nei M (1978) Estimation of average heterozygocity and genetic distance from a small number of individuals. Genetics 89:583–590

Nielsen R (2005) Molecular signatures of natural selection. Annu Rev Genet 39:197–218

Pañeda A, Rodríguez-Suárez C, Campa A, Ferreira JJ, Giraldez R (2008) Molecular markers linked to the *fin* gene controlling determinate growth habit in common bean. Euphytica 162:241–248

Papa R, Acosta J, Delgado-Salinas A, Gepts P (2005) A genome-wide analysis of differentiation between wild and domesticated *Phaseolus vulgaris* from Mesoamerica. Theor Appl Genet 111:1147–1158

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

Pastor-Corrales MA (1991) Estandarización de variedades diferenciales y de designación de razas de *Colletotrichum 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

Peakall R, Smouse PE (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Piergiovanni A, Taranto G, Losavio FP, Pignone D (2006) Common bean (*Phaseolus vulgaris* L.) landraces from Abruzzo and Lazio regions (Central Italy). Genet Resour Crop Evol 53:313–322

Pritchard JK, Wen W (2004) Documentation for the STRUCTURE software Version 2. Chicago. http://www.pritch.bsd.uchicagoedu/software/structure2_1.html

Pritchard JK, Stephens M, Donnelly P (2000) Inference of landrace structure using multilocus genotype data. Genetics 155:945–959

Rick CM, Holle M (1990) Andean *Lycopersicon esculentum* var. cerasiformie. Genetic variation and its evolutionary significance. Econ Bot 44:69–78

Rodiño P, Santalla M, González AM, De Ron AM, Singh SP (2006) Novel genetic variation in common bean from the Iberian Peninsula. Crop Sci 46:2540–2546

Rosales-Serna R, Hernandez-Delgado S, Gonzalez-Paz M, Acosta-Gallegos JA, Mayek-Perez N (2005) Genetic relationships and diversity revealed by AFLP markers in Mexican common bean bred cultivars. Crop Sci 45:1951–1957

Rossi M, Bitocchi E, Bellucci E, Nanni L, Rau D, Attene G, Papa R (2009) Linkage disequilibrium and population structure in wild and domesticated populations of *Phaseolus vulgaris* L. Evol Appl. doi:10.1111/j.1752-4571.2009.00082.x

Sánchez E, Sifres A, Casañas F, Nuez F (2008) The endangered future of organoleptically prestigious European landraces: Ganxet bean (*Phaseolus vulgaris* L.) as an example of a crop originating in the Americas. Genet Resour Crop Evol 55:45–52

Santalla M, Rodiño AP, De Ron AM (2002) Allozyme evidence supporting southwestern Europe as a secondary center of genetic diversity for common bean. Theor Appl Genet 104:934–944

Schlötterer C (2003) Hitchhiking mapping—functional genomics from the population genetics perspective. Trends Genet 19:32–38

Sicard D, Nanni L, Porfiri O, Bulfon D, Papa R (2005) Genetic diversity of *Phaseolus vulgaris* L. and *Phaseolus coccineus* L. landraces in central Italy. Plant Breed 124:464–472

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

Singh SP, Gutiérrez JA (1990) Effect of plant density on selection for seed yield in two landrace 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 traits. Crop Sci 31:23–29

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

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

Slatkin M, Barton N (1989a) A comparison of three indirect methods for estimating average levels of gene flow. Evolution 43:1349–1368

Slatkin M, Barton NH (1989b) A comparison of three indirect methods for estimating the average level of gene flow. Evolution 43:1349–1368

Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive landrace divergence. Mol Ecol 14:671–688

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

CIAT (Centro Internacional de Agricultura Tropical) (1987) Standard Voysest O. 2000. Mejoramiento genético del frijol (*Phaseolus vulgaris* system for the evaluation of bean germplasm. Van Schoohoven L.): Legado de variedades de América Latina 1930–1999. Pastor- Corrales (compilers). CIAT. Cali, Colombia

van Schoonhoven A, Pastor-Corrales MA (1987) Standard system for the evaluation of bean germplasm. CIAT, Cali, Colombia

Vasemägi A, Nilsson J, Primmer CR (2005) Expressed sequence tag (EST) linked microsatellites as a source of gene associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar* L.). Mol Biol Evol 22:1067–1076

Vavilov NI (1951) The origin, variation, immunity and breeding of cultivated plants. Chron Bot 13:1–366 Translated by K. Start

Vigouroux Y, MacMullen M, Hittinger CT, Houchins K, Schulz L, Kresovich S, Matsuoka Y, Doebley J (2002) Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. Proc Natl Acad Sci USA 99:9650–9655

Vitalis R, Dawson K, Boursot P (2001) Interpretation of variation across marker loci as evidence of selection. Genetics 158:1811–1823

Vitalis R, Dawson K, Boursot P, Belkhir K (2003) DetSel 1.0: a computer program to detect markers responding to selection. J Hered 94:429–431

Vitte C, Ishii T, Lamy F, Brar D, Panaud O (2004) Genomic paleontology provides evidence for two distinct origins of Asian rice (*Oryza sativa* L.). Mol Genet Genom 272:504–511

Wang L, Guan R, Zhangxiong L, Chang R, Qiu L (2006) Genetic Diversity of Chinese Cultivated Soybean Revealed by SSR Markers. Crop Sci 46:1032–1038

Weir BS, Cockerman C (1984) Estimating F-statistics for the analysis of landrace structure. Evolution 38:1358–1370

White JW, Singh SP, Pino C, Rios MJ, Buddenhagen I (1992) Effects of seed size and photoperiod response on crop growth and yield of common bean. Field Crops Res 23:159–175

Xu P, Tao DY, Hu FY, Zhou JW, Deng XN (2005) Interspecific hybridization of cultivated rice for breeding japonica rice in Yunnan province. Chin J Rice Sci 19(1):41–46

Yaish MWF, Pérez De La Vega M (2003) Isolation of $(GA)_n$ microsatellite sequences and description of a predicted MADS-box sequence isolated from common bean (*Phaseolus vulgaris* **L.**). Genet Mol Biol 26:337–342

Yu K, Park J, 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

Zizumbo-Villarreal D, Colunga-García Marín P, De la Cruz E, Payró Delgado-Valerio P, Gepts P (2005) Landrace structure and evolutionary dynamics of wild–weedy–domesticated complexes of common bean in a Mesoamerican region. Crop Sci 45:1073–1083