ORIGINAL PAPER

R. Papa · J. Acosta · A. Delgado-Salinas · P. Gepts

A genome-wide analysis of differentiation between wild and domesticated *Phaseolus vulgaris* from Mesoamerica

Abstract Lack of introgression or divergent selection may be responsible for the maintenance of phenotypic differences between sympatric populations of crops and their wild progenitors. To distinguish between these hypotheses, amplified fragment length polymorphism markers were located on a molecular linkage map of *Phaseolus vulgaris* relative to genes for the domestication syndrome and other traits. Diversity for these same markers was then analyzed in two samples of wild and domesticated populations from Mesoamerica. Differentiation between wild and domesticated populations was significantly higher in parapatric and allopatric populations compared to sympatric populations. It was also significantly higher near genes for domestication compared to those away from these genes. Concurrently, the differences in genetic diversity between wild and domesticated populations were strongest around such genes. These data suggest that selection in the presence of introgression appears to be a major evolutionary factor maintaining the identity of wild and domesticated populations in sympatric situations. Furthermore, alleles from domesticated populations appear to have displaced alleles in sympatric wild populations, thus leading to a reduction in genetic diversity in such populations. These results also provide a possible experimental framework for assessing the long-term risk of transgene escape and the targeting of transgenes inside the genome to minimize the survival of these transgenes into wild populations following introduction by gene flow.

This article is dedicated to the memory of Epimaki M. K. Koinange.

R. Papa · P. Gepts (✉)
Department of Plant Sciences, Section of Crop and Ecosystem Sciences, University of California, Mailstop 1, 1 Shields Avenue, Davis, CA 95616-8780, USA
E-mail: plgepts@ucdavis.edu
Tel.: +1-530-7527743
Fax: +1-530-7524361

R. Papa
Dipartimento di Scienze degli Alimenti, Facoltà di Agraria, Universitá Politecnica delle Marche, Via Brecce Bianche, 60131 Ancona, Italy

J. Acosta
Instituto Nacional de Investigaciones Forestales y Agropecuarias, Campo Experimental del Bajío, Apartado Postal 112, Guanajuato, CP 38000 Celaya, Mexico

A. Delgado-Salinas
Instituto de Biología, Universidad Nacional Autónoma de México, Apartado Postal 70-233, 04510 México D.F., Mexico

Introduction

One of the single most important and generalized features of plant domestication is the reduction in genetic diversity that has characterized crop gene pools, not only during the initial domestication phase itself but also subsequently during dispersal from centers of domestication and further selection (Gepts 2004). This reduction is caused by both demographic events (for example, a bottleneck due to a reduction in population size) and selection (for example, for adaptation to a cultivated environment) (Vigouroux et al. 2002; Tenaillon et al. 2004). As posited by Cavalli-Sforza (1966), demographic events and selection in populations affect genetic diversity in different ways. While demographic events affect the genetic diversity across an entire genome, selection will affect only specific genes or adjacent genes. Hence, patterns of genetic differentiation, assessed with presumably neutral markers, reflect demographic or historical parameters, such as effective population size, population bottlenecks or expansions, reproductive system, migration, and time since divergence. In addition, they may also be affected by selection on adjacent, linked loci. For example, stabilizing or balancing selection can lead to elevated levels of variation at closely linked neutral loci. Directional selection for local adaptation will increase differentiation at those loci under selection leading to high F_{ST} values for closely linked neutral loci (Charlesworth et al. 1997). Neutral

markers can be subjected to genetic hitchhiking or a selective sweep due to selection for advantageous mutations. They can also be subject to background selection due to selection against deleterious mutations. Both types of selection will reduce genetic diversity around selected locus and increase population differentiation.

Several studies have been performed that show variability for population differentiation among loci (Merilä and Crnokrak 2001; McKay and Latta 2002). In reviewing data from 29 species, McKay and Latta (2002) found that in 24 of 29 species Q_{ST} (an analogous parameter to F_{ST} but for QTL data) averaged over traits was higher than F_{ST} averaged over marker loci, suggesting an important role for selection. However, with few exceptions, these putatively selected loci have not been related to a demonstrably adaptive locus. The domestication of plants and animals is a situation in which some of the morphological and physiological traits under selection—the so-called domestication syndrome (Gepts 2004)—are well known in major crops. Domestication includes primarily a selection for reduced seed dispersal and dormancy, a more compact growth habit, reduced sensitivity to daylength and tendency towards increased selfing or vegetative propagation, and an increase in diversity and size of the harvested parts. Genes controlling these traits have been mapped in several crops, including maize (Doebley et al. 1990; Doebley and Stec 1991), common bean (Koinange et al. 1996), pearl millet (Poncet et al. 1998, 2000, 2002), rice (Xiong et al. 1999), and sunflower (Burke et al. 2002). While the genetic architecture of domestication differs in the details among the crops analyzed, information about the location on the genetic map of genes that have been selected during domestication provides an excellent model to study the effect of selection on the structure of genetic diversity along the genome. We predicted that domestication loci or closely linked, neutral loci would show differences in frequencies and differentiation between wild and domesticated populations compared with other regions of the genome that are not involved in domestication. Conversely, loci or genome regions with significantly different gene frequencies may have been potential target regions for selection operating during domestication.

A potential confounding factor is the presence of gene flow between wild and domesticated populations. We expected that if gene flow between wild and domesticated populations were rare, the differentiation between these two types of populations would trace back to the initial domestication phase, which represents a combination of two evolutionary processes: (1) selection for specific traits belonging to the domestication syndrome; and (2) genetic drift for the rest of the genome due to the bottleneck of domestication and isolation after domestication. Both evolutionary processes will lead to significant differentiation throughout the genome between wild and domesticated types. In contrast, if gene flow plays a significant role, then selection, which maintains the wild and domesticated phenotypes, may only lead to divergence at loci responsible for phenotypic differences between wild and domesticated types and at linked loci. Differentiation will be largest around domestication genes and smaller away from these domestication genes depending on the level of gene flow. Thus, in such a system selection acts to prevent the invasion of certain genome regions by domesticated alleles in wild types and vice versa. Away from these domestication loci, gene flow could lead to replacement of native alleles. Of particular concern, in this context, might be the displacement of wild alleles given the steady genetic erosion to which wild populations have been subjected to. Here, we report on a genome-wide analysis of differentiation between wild and domesticated common bean (*Phaseolus vulgaris* L.), in which we identify a statistically significant higher differentiation between these populations around domestication genes in comparison with genomic regions away from domestication loci.

Materials and methods

Plant material and molecular analysis

Two sets of plant materials were analyzed. First, samples were collected on a single plant basis during an exploration conducted in December 1996 in the Mexican state of Chiapas from 12 wild and 10 domesticated populations, representing 221 individuals (Fig. 1; Table 1). The area of domestication of *P. vulgaris* in Mesoamerica is tentatively located in Central North Mexico, presumably in the current states of Jalisco and Guanajuato and far from Chiapas (Gepts 1988). Moreover, wild populations from Chiapas are highly differentiated from Central North Mexico populations (Papa and Gepts 2003). For both reasons, Chiapas was a well-suited location in order to assess the role of gene flow after domestication because the confounding effect of shared ancestry in the assessment of gene flow would be reduced. Based on farmer information, all domesticated materials were traditional landraces. Morphological traits and molecular data (Papa and Gepts 2003) showed that the wild populations were neither weedy nor escapes from cultivation. There was a marked phenotypic differentiation between wild and domesticated types, regardless of where the wild plants grew (within or around fields). The 22 Chiapas populations collected in Chiapas represented three levels of spatial arrangement. Allopatry was represented by wild and domesticated populations growing in three different regions of the state (Tuxtla, Teopisca, and Las Rosas; Fig. 1a) or at a distance larger than 1 km within a single region (i.e., Teopisca: D, C, F vs E, I, J; Fig. 1a, b). Parapatry was assumed to include wild and domesticated plants found in spatially distinct populations distant from a few meters to 1 km (e.g., wild and domesticated populations in site J_N; Fig. 1b). Sympatry was assumed to consist of

Fig. 1 Geographic distribution of wild and domesticated populations of *P. vulgaris* from Chiapas. **a** Main collection areas in Chiapas: Tuxtla (sites A, G, and H), Teopisca (sites C–F, I, and J), and Las Rosas (M–O). **b** Spatial arrangement of populations in sites E, I, and J of the Río Blanco site within the Teopisca area
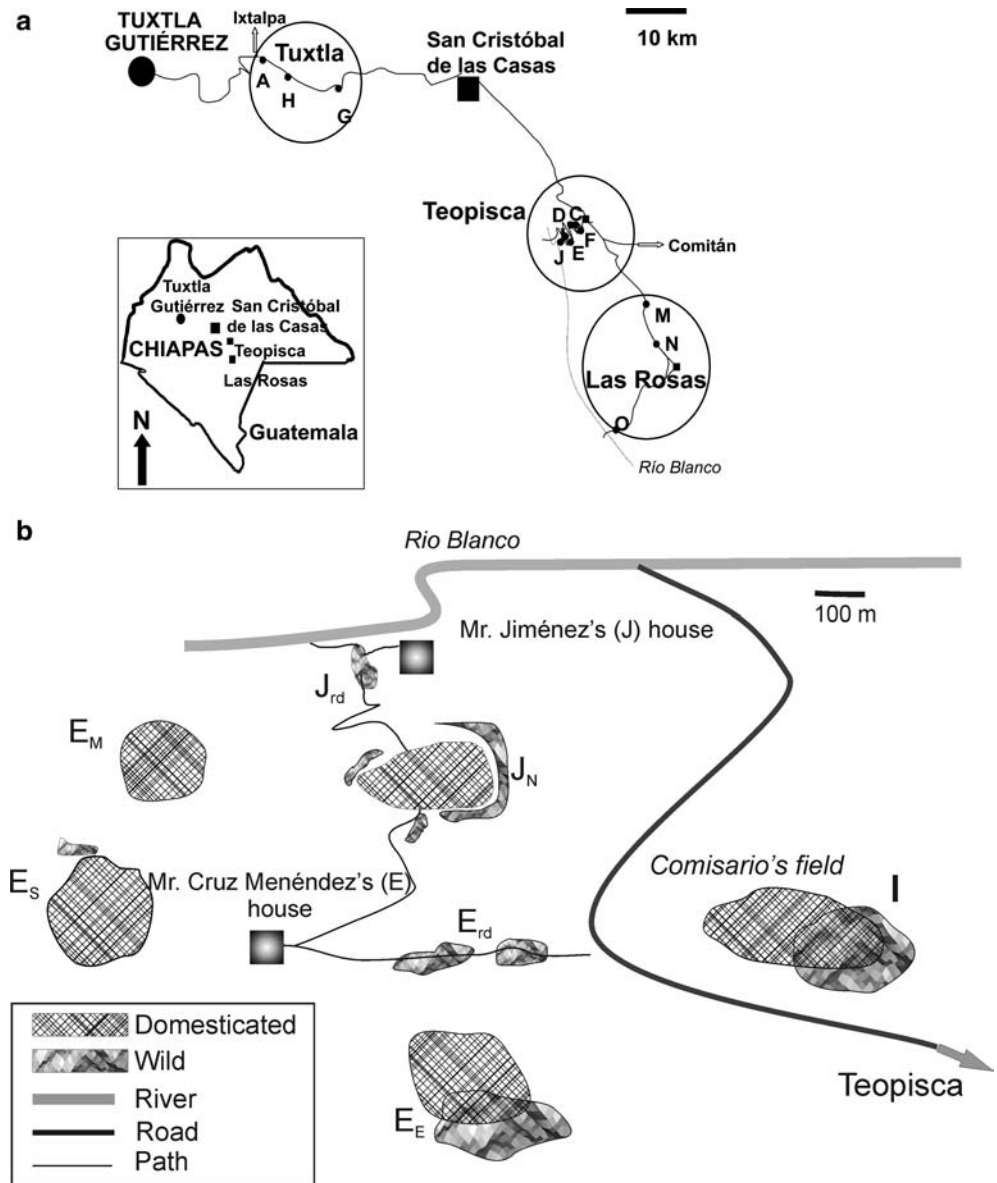
sites where portions of the wild plants were growing in fields of domesticated plants (e.g., populations E_E and I, Fig. 1b). Farmer information indicates that the local geographic distribution of the different domesticated populations and their place in crop rotations did not change appreciably over time, hence the spatial relationships between wild and domesticated populations remained similar over time. Both cluster and spatial autocorrelation analyses presented in Papa and Gepts (2003) showed that two domesticated populations were genetically distinct from all the other domesticated populations from Chiapas probably because of their phenology (i.e., determinate type vs indeterminate) and agronomic system (i.e., pure stand vs association with maize plants). Because they probably represent recent, non-traditional introductions into the region, they were not included in the present study. Thus, overall 20 of the 22 collected populations were used.

Second, analyses were also conducted in a geographically broader sample representative of the Mesoamerican gene pool and consisted of 25 domesticated and 61 wild genotypes from Mexico and Central America. A complete list and description of the accessions analyzed are reported in Papa and Gepts (2003). AFLPs were analyzed in these two sets of materials according to published procedures (Vos et al. 1995). Four *Eco*RI/ *Mse*I primer combinations were used with the following selective bases (5′-3′/5′-3′): AAC/AGG, ACA/ACA, ACC/ATG, AGT/AGA, and AGC/ATG (the latter used only in the Mesoamerican sample).

Linkage mapping

AFLPs were mapped among 56 lines of the F8 recombinant inbred population BAT93 × Jalo EEP588

Table 1 Common bean populations from Chiapas in areas with sympatric wild and domesticated beans

Locations			Materials collected			
			Wild		Domesticated	
Area	Altitude (m, a.s.l.)	Sites	Population name	Number of individuals analyzed	Population name	Number of individuals analyzed
Tuxtla	1,300	A	CHWA	9	CHCA	2
	1,360	H	CHWH	12		
	1,450	G	CHWG	14		
Teopisca	1,650	F			CHCF	8
	1,450	D			CHCD	8
	1,200	I	CHWI	12	CHCI	22
	1,200–1,050	E South	CHWE$_S$	9	CHCE$_S$	20
		E East	CHWE$_E$	18	CHCE$_E$	18
		E Road	CHWE$_{RD}$	8		
		J North	CHWJ$_N$	13	CHCJ$_N$	12
		J Road	CHWEJ$_{RD}$	6		
Las Rosas	1,550	M	CHWM	3	CHCM	8
	1,430	N	CHWN	9		
	1,180	O	CHWO	10		
Total				123		98

according to standard mapping procedures (Freyre et al. 1998). The core linkage map previously established in this population contained some 560 markers distributed in 11 linkage groups. Linkage map distances in Kosambi units were determined between 51 mapped AFLPs and previously mapped linked genes or QTLs for phenotypic traits (Geffroy et al. 2000; Gepts 1999; Koinange et al. 1996; Nodari et al. 1993). Linkage distances between AFLP markers and QTLs were calculated to the adjacent markers with the highest LOD score. Although the confidence interval of the location of QTLs in a linkage map is often larger than that for a major gene, the location of the maximum LOD score value is a sufficiently precise approximation of the actual location of the QTL in our experience. In common bean, this is supported by prior research showing the co-segregation of the fin locus for determinacy and a QTL for the number of nodes on the main stem (Koinange et al. 1996), of a QTL for resistance to white mold (*Sclerotinia sclerotiorum*) and growth habit genes (Miklas et al. 2001), and of major genes and QTLs for anthracnose (*Colletotrichum lindemuthianum*) resistance with resistance gene analogs (Geffroy et al. 2000).

Data analysis

Data were analyzed assuming a haploid genome (i.e., complete homozygosity) because of the predominantly selfing mating system of *P. vulgaris* (over 98% in most studies: Ibarra-Pérez et al. 1997). High levels of homozygosity have been observed in both wild and domesticated beans through the use of codominant markers such as phaseolin (Gepts et al. 1986) and allozymes (Koenig and Gepts 1989; Singh et al. 1991). Microsatellite marker analyses of original seeds collected in wild populations from the state of Morelos (Mexico) showed a F_{IT} value of 0.97, further confirming the high level of homozygosity in wild beans as well (D. Sicard and R. Papa, unpublished results). Genetic differentiation was assessed in the Chiapas populations using the F_{ST} parameter calculated according to the weighted average F-statistics over loci (Weir and Cockerham 1984). For neutral markers, differentiation between populations can arise by drift alone; hence, homogeneous values of F_{ST} are expected across the genome. Loci under selection are poor indicators of gene flow because, in addition to drift, they reflect the action of selection and, compared with neutral loci, will show higher values of F_{ST} if selection is heterogeneous across population and lower values in presence of homogeneous selection. Molecular markers such as AFLP are presumed to be neutral but, particularly in selfing species such as *P. vulgaris*, hitchhiking and background selection may strongly influence their behavior. Thus, to estimate the level of gene flow between the wild and domesticated populations in the Chiapas population using the F_{ST} statistic, 20 markers linked (LOD score > 3.0; Freyre et al. 1998) to genes previously identified in *P. vulgaris* (Geffroy et al. 2000; Gepts 1999; Koinange et al. 1996; Nodari et al. 1993) were excluded from the analysis. In addition, three unmapped markers that were significantly correlated ($P < 0.01$) with the altitude of collection sites were also excluded from this analysis. Overall, 23 markers out of 101 were thus excluded. Thus, 78 AFLP markers were used in the analysis. Gene diversity (H; Nei and Li 1979) was calculated for each population. The significance of F_{ST} estimates was calculated using a non-parametric permutation approach consisting in permuting haplotypes among populations (Excoffier et al. 1992) and implemented in the software Arlequin ver. 2 (http://anthropologie.unige.ch/arlequin). Pairwise F_{ST} values are reported in Table 2. The Wilcoxon nonparametric test was used to test the

Table 2 F_{ST} and H values for AFLP markers as a function of linkage map location for the Mesoamerica and Chiapas samples

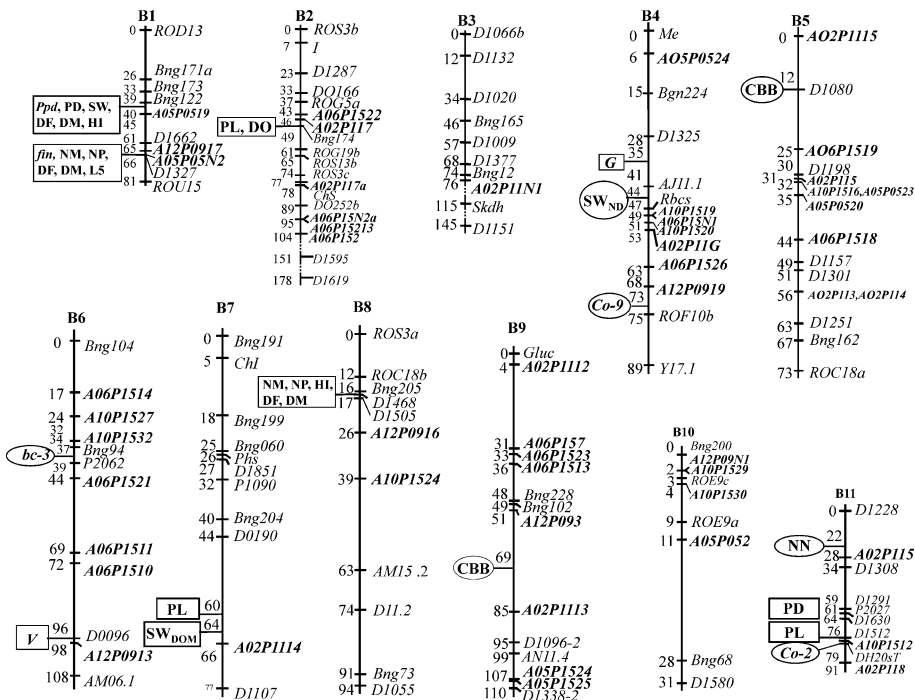| Linkage group | AFLP marker[a] | Nearest linked trait | | Mesoamerica | | | | Chiapas | | | |
		Distance (cM)[b]	Traits[c]	F_{ST}	H_{Wild}	H_{Dom}	H_{Total}	F_{ST}	H_{Wild}	H_{Dom}	H_{Total}
DOM markers (linked to genes for domestication)											
B01	A05P0519	5	*Ppd*, PD, SW$_{DOM}$, DF, DM, HI	0.23	0.51	0.28	0.46	0.06	0.48	0.49	0.50
B01	A12P0917	1	*fin*, NM, NP, DF, DM, L5	0.48	0.48	0.08	0.45	0.58	0.45	0.07	0.48
B01	A05P05N2	0	*fin*, NM, NP, DF, DM, L5	0.73	0.25	0.16	0.48	0.89	0.21	0.05	0.50
B02	A06P1522	6	PL, DO	0.27	0.50	0.28	0.48	0.76	0.24	0.14	0.51
B02	A02P117	3	PL, DO	0.35	0.44	0.33	0.50	0.50	0.50	0.02	0.43
B04	AO5P0524	29	*G*	0.10	0.51	0.42	0.49	0.10	0.29	0.20	0.25
B06	A12P0913	3	*V*	0.12	0.32	0.00	0.18	–	–	–	–
B07	A02P1114	2	SW$_{DOM}$, PL	0.20	0.28	0.52	0.45	0.00	0.11	0.07	0.09
B08	A12P0916	9	NM, NP, HI, DF, DM	0.32	0.51	0.08	0.39	0.36	0.46	0.30	0.50
B08	A10P1524	24	NM, NP, HI, DF, DM	0.33	0.50	0.00	0.34	–	–	–	–
B11	A10P1512	1	PL, PD	0.07	0.23	0.00	0.13	–	–	–	–
Averages D markers		7		0.29	0.41	0.20	0.40	0.41	0.34	0.17	0.41
ND markers (linked to non-domestication genes)											
B04	A06P15N1	4	SW$_{ND}$	0.12	0.49	0.45	0.51	0.00	0.37	0.29	0.31
B04	A10P1519	5	SW$_{ND}$	0.31	0.26	0.52	0.46	–	–	–	–
B04	A06P15N1	5	SW$_{ND}$	0.37	0.21	0.52	0.45	–	–	–	–
B04	A02P11G	9	SW$_{ND}$	0.00	0.07	0.00	0.03	–	–	–	–
B04	A06P1526	10	*Co-9*	0.01	0.37	0.22	0.30	0.46	0.44	0.23	0.50
B04	A12P0919	5	*Co-9*	0.32	0.51	0.08	0.39	0.62	0.45	0.02	0.47
B05	AO2P1115	12	CBB	0.00	0.49	0.50	0.49	0.06	0.29	0.09	0.21
B05	AO6P1519	13	CBB	0.00	0.13	0.08	0.10	0.00	0.02	0.05	0.03
B05	A02P115	19	CBB	0.02	0.50	0.50	0.51	0.11	0.47	0.47	0.51
B05	A10P1516	20	CBB	0.08	0.26	0.46	0.37	–	–	–	–
B05	A05P0523	20	CBB	0.00	0.03	0.00	0.02	0.00	0.02	0.00	0.01
B05	A05P0520	23	CBB	0.33	0.51	0.15	0.44	0.76	0.33	0.02	0.50
B06	A06P1514	20	*bc-3*	0.17	0.48	0.42	0.51	0.00	0.18	0.20	0.19
B06	A10P1527	13	*bc-3*	0.08	0.49	0.49	0.51	–	–	–	–
B06	A10P1532	5	*bc-3*	0.01	0.38	0.49	0.44	–	–	–	–
B06	A06P1521	7	*bc-3*	0.00	0.03	0.00	0.02	–	–	–	–
B09	A02P1113	16	CBB	0.81	0.21	0.08	0.51	0.90	0.12	0.05	0.50
B11	A02P115	6	NN	0.00	0.03	0.00	0.02	0.00	0.02	0.00	0.01
B11	A02P118	12	*Co-2*	0.00	0.35	0.45	0.40	0.06	0.35	0.48	0.42
Averages ND markers		12		0.14	0.31	0.28	0.34	0.25	0.26	0.16	0.31
UN markers (unlinked to known genes)											
B02	A02P117a	38[c]		0.01	0.21	0.08	0.15	0.00	0.00	0.02	0.01
B02	A06P15N2a	45		0.33	0.15	0.51	0.31	0.02	0.02	0.09	0.05
B02	A06P15213	45		0.14	0.49	0.22	0.39	0.00	0.03	0.02	0.03
B02	A06P152	50		–	–	–	–	0.01	0.03	0.00	0.02
B03	A02P11N1	40		0.10	0.28	0.00	0.16	0.04	0.09	0.00	0.05
B05	A06P1518	33		0.41	0.16	0.52	0.43	0.00	0.19	0.27	0.23
B05	AO2P113	44		0.07	0.23	0.00	0.13	–	–	–	–
B05	AO2P114	44		0.06	0.21	0.00	0.11	–	–	–	–
B06	A06P1511	38		0.37	0.16	0.52	0.41	0.13	0.16	0.43	0.30
B06	A06P1510	35		0.17	0.35	0.52	0.48	0.15	0.12	0.41	0.27
B09	A02P1112	50		0.02	0.50	0.42	0.47	0.06	0.46	0.31	0.41
B09	A06P157	50		0.00	0.21	0.15	0.18	0.00	0.02	0.02	0.02
B09	A06P1523	50		0.00	0.10	0.08	0.09	0.00	0.19	0.27	0.23
B09	A06P1513	50		0.04	0.16	0.00	0.08	0.00	0.00	0.02	0.01
B09	A12P093	40		0.02	0.23	0.38	0.31	0.05	0.24	0.07	0.17
B09	A05P1524	38		0.32	0.51	0.08	0.39	0.48	0.50	0.05	0.43
B09	A05P1525	38		0.20	0.47	0.08	0.33	0.48	0.50	0.05	0.44
B10	A10P1529	50		0.04	0.03	0.16	0.10	–	–	–	–
B10	A12P09N1	50		0.31	0.51	0.16	0.49	0.14	0.36	0.03	0.21
B10	A10P1530	50		0.00	0.03	0.00	0.02	–	–	–	–
B10	A05P052	50		0.05	0.45	0.52	0.50	0.00	0.40	0.39	0.40
Averages UN markers		42		0.13	0.27	0.22	0.28	0.09	0.19	0.14	0.19
General averages		23		0.17	0.32	0.24	0.33	0.21	0.25	0.15	0.28

[a]For map location, see Fig. 3; [b]Distance to closest known gene or QTL; [c]Major genes (italics) or QTLs (upper case): *bc-3* bean common mosaic virus resistance; *CBB* common bacterial blight resistance; *Co-2* and *Co-7* anthracnose resistance; *DF* days to flowering; *DM* days to maturity; *DO* dormancy; *fin* determinacy; *G* seed color; *HI* harvest index; *L5* length of the fifth internode; *NM* number of nodes on the main stem; *NN* Rhizobium nodule number; *NP* number of pods; *PL* pod length; *Ppd* photoperiod sensitivity; *PD* photoperiod sensitivity; SW$_{DOM}$ seed weight, identified in cross with wild bean; SW$_{ND}$ seed weight, identified in cross between cultivars; *V* flower and seed color (Geffroy et al. 2000; Gepts 1999; Koinange et al. 1996; Nodari et al. 1993)

significance ($P < 0.05$) of differences between F_{ST} and H values obtained for different groups of populations (Sokal and Rohlf 1995).

The effect of selection was estimated by calculating F_{ST} and H for all the mapped markers found to be polymorphic in the Chiapas or Mesoamerican samples, including those that they were linked to genes or were correlated with altitude. F_{ST} and H values for individual loci as a function of linkage map location are reported in Table 2. A Wilcoxon nonparametric test was used to test the significance ($P < 0.05$) of differences between F_{ST} and H values obtained for different groups of markers. Kendall's nonparametric correlation coefficient (tau) was used to test association between the estimates of the same parameter (F_{ST} or H) obtained in the Chiapas and Mesoamerica samples as well as to test the association between F_{ST} and cM distances between markers and major genes (Sokal and Rohlf 1995).

Fig. 2 Molecular linkage map of common bean. To the right of each linkage group are AFLP markers (name in *bold type and starting with A*) and previously mapped framework markers (Freyre et al. 1998). To the left are phenotypic traits mapping near AFLP markers: domestication traits (*rectangles*) and other traits (*ovals*). Genetic distances are in Kosambi map units. Major genes (*italics*) or QTLs (*upper case*): *bc-3*, bean common mosaic virus resistance; *CBB*, common bacterial blight resistance; *Co-2* and *Co-7*, anthracnose resistance; *DF*, days to flowering; *DM*, days to maturity; *DO*, dormancy; *fin*, determinacy; *G*, seed color; *HI*, harvest index; *L5*, length of the fifth internode; *NM*, number of nodes on the main stem; *NN*, *Rhizobium* nodule number; *NP*, number of pods; *PL*, pod length; *Ppd*, photoperiod sensitivity; *PD*, photoperiod sensitivity; SW_{DOM} seed weight, identified in cross with wild bean; SW_{ND} seed weight, identified in cross between cultivars; *V*, flower and seed color (Geffroy et al. 2000; Gepts 1999; Koinange et al. 1996; Nodari et al. 1993)

Results

The genetic diversity and population differentiation between wild and domesticated populations for single AFLP markers were analyzed relative to their map positions and linkage distance from known genes. We used two sets of accessions sampled at different geographical scales, one originating from an exploration of populations conducted in Chiapas and another obtained from Phaseolus gene banks representing a geographically broader sample of the Mesoamerican gene pool and consisting of 25 domesticated and 61 wild genotypes from Mexico and Central America. In the latter sample, an additional AFLP primer combination compared to the Mesoamerican sample (see Materials and methods) was used to analyze genetic diversity. In both the Chiapas and Mesoamerican samples, all individuals were grouped into either a wild or domesticated population. Genetic differentiation was calculated between these two pooled populations. Both parameters (H and F_{ST}) were highly correlated between the Chiapas and Mesoamerica samples based on tau, Kendall's nonparametric correlation coefficient. The respective values of tau for the different H parameters were, for wild beans 0.55 ($P = 0.0001$), for domesticated beans 0.47 ($P = 0.0001$), and for the entire population 0.36 ($P = 0.0023$). The tau value for F_{ST} was 0.46 ($P = 0.0002$).

Linkage mapping of AFLP markers

The AFLP markers were located on all linkage groups, from one marker on B3 and B7 to nine markers on B5 with an average of five markers per linkage group

(Fig. 2). AFLPs were classified into three classes depending on their linkage relationships with genes related to domestication and other traits (Table 2): (1) UN, 19 and 20 markers unlinked (>30 cM) to any known gene or QTL for the Chiapas and Mesoamerica samples, respectively; (2) ND, 12 and 19 markers, respectively, linked to a gene or QTL identified in domesticated × domesticated crosses but presumably not involved in the domestication syndrome (Geffroy et al. 2000; Gepts 1999; Nodari et al. 1993); and (3) D, 8 and 11 markers, respectively linked to a gene or QTL involved in the domestication syndrome of *P. vulgaris* (Koinange et al. 1996).

Linkage map-based distribution of population differentiation (F_{ST}) and genetic diversity (H)

All the analyses conducted in the two samples individually as well as in the pooled sample (Table 3) showed that, according to a Wilcoxon nonparametric test, AFLPs unlinked to any known gene or QTL (UN: $F_{ST} = 0.13$ Mesoamerica; 0.09 Chiapas) were less differentiated between wild and domesticated population than AFLPs linked to genes for the domestication syndrome (D: $F_{ST} = 0.29$ Mesoamerica; 0.41 Chiapas). While F_{ST} estimates obtained for ND markers ($F_{ST} = 0.14$ Mesoamerica; 0.25 Chiapas) were never significantly different from those obtained for UN markers, with the exception of the Chiapas sample, they were significantly lower than those obtained for D markers.

Genetic diversity was estimated by Nei's unbiased estimator (Nei 1987) and by calculating genotypic richness (the number of different genotypes across population and marker classes). Because the results for the comparisons between wild and domesticated populations for the D, ND, and UN markers, as well as the pooled marker group, were similar, only the Nei diversity results are presented here (Table 4). When all markers were considered, the genetic diversity of the domesticated population was always significantly lower than that of the wild populations with the exception of the Mesoamerican sample. A significantly higher genetic diversity for wild populations was also detected for D markers in all the samples. In contrast, for both ND and UN markers, differences in genetic diversities between wild and domesticated populations were never significant. When the genetic diversity of the domesticated populations (H_D) was compared among the three marker classes (Table 3), none of the estimates showed significant differences neither in the Chiapas nor the Mesoamerica samples. However, the genetic diversity of wild populations (H_{Wild}) was significantly higher for D markers than for UN markers for the pooled and the Mesoamerican samples, also after Bonferroni correction, but not for the Chiapas samples (although the difference was statistically nearly significant: $P = 0.054$). The other comparisons (D vs ND and ND vs UN) were never significant. The total genetic diversity (H_T) was always higher for D markers compared to UN markers and, in the pooled sample, was higher for ND markers than for UN markers.

Association between marker linkage distances and genetic parameters

The association of linkage distances between markers and domestication genes and both F_{ST} and genetic diversity (H_{Wild}, H_{Dom} and H_{Total}) was also investigated.

Table 3 Statistical analyses of population genetic parameters assessed with different classes of AFLP markers

		F_{ST}	H_{Wild}	H_{Dom}	H_{Total}
Pooled sample: No. of markers: D (19) ND (31) UN (37)					
D vs UN	χ^2	13.129	9.5735	0.04	13.23
	$P > \chi^2$	*0.0003*[b]	*0.0020*	0.8416	*0.0003*
D vs ND	χ^2	7.3592	3.5771	0.4253	0.7433
	$P > \chi^2$	*0.0067*	0.0586	0.5143	0.3886
ND vs UN	χ^2	0.001	1.1246	0.6976	4.8962
	$P > \chi^2$	0.9750	0.2889	0.4036	**0.0269**
Mexico sample: No. of markers: D (11) ND (19) UN (20)					
D vs UN	χ^2	6.2529	6.2872	0.0632	4.2734
	$P > \chi^2$	*0.0124*	*0.0122*	0.8015	**0.0387**
D vs ND	χ^2	6.1004	2.8392	1.1799	0.1194
	$P > \chi^2$	*0.0135*	0.0920	0.2774	0.7297
ND vs UN	χ^2	0.8234	0.3675	0.2763	1.7136
	$P > \chi^2$	0.3642	0.5444	0.5991	0.1905
Chiapas sample: No. of markers: D (8) ND (12) UN (17)					
D vs UN	χ^2	6.1912	3.71	0.7725	8.6848
	$P > \chi^2$	*0.0128*	0.0541	0.3794	*0.0032*
D vs ND	χ^2	1.3808	1.5307	0.4336	1.3776
	$P > \chi^2$	0.2400	0.2160	0.5102	0.2405
ND vs UN	χ^2	0.575	0.474	0.0244	2.2773
	$P > \chi^2$	0.4483	0.4912	0.8760	0.1313

[a] Wilcoxon/Kruskal–Wallis tests (rank sums) one-way test. χ^2 approximation (1 df);
[b] Significant differences ($P < 0.05$) before (bold) and after (italics) Bonferroni correction

Table 4 Comparison of genetic diversity (H) in wild and domesticated populations assessed with markers in different locations of the common bean genome

Markers	Pooled sample		Mesoamerica		Chiapas	
	χ^2	$P > \chi^{2a}$	χ^2	$P > \chi^2$	χ^2	$P > \chi^2$
All	9.31	0.0023	3.54	0.0597	4.70	0.0302
D	10.62	0.0011	5.79	0.0161	3.99	0.0458
ND	1.00	0.3169	0.06	0.8035	1.49	0.2229
UN	2.28	0.1308	1.21	0.2717	0.69	0.4065

[a]Wilcoxon/Kruskal–Wallis tests (rank sums) one-way test. χ^2 approximation (df = 1)

AFLP markers were classified into two groups: (1) ND markers and UN markers placed on linkage groups where ND loci were identified; and (2) D markers and UN markers placed on linkage groups where D loci were identified. In both cases, the shortest genetic (cM) distance between a marker and a QTL or gene was used. The maximum value of linkage distance considered was 50 cM even for markers that were at located at a higher linkage distance from genes. No significant correlation was found between map distances from ND markers to QTLs or genes, on the one hand, and F_{ST}, H_{Wild}, H_{Dom} and H_{Total}, on the other, regardless of the geographic sample used (Chiapas or Mesoamerica). When D markers were considered, a significant negative correlation was found between map distance and F_{ST}, H_{Wild}, and H_{Total} in the Chiapas sample and in the pooled Chiapas and Mesoamerican samples (Table 5). In the Mesoamerican sample, no significant correlation was found. Overall, these data suggest that the observed pattern of F_{ST} variation is related to the effect of genes involved in the domestication syndrome undergoing bidirectional selection for alternative alleles in the wild and domesticated environment. The same results may suggest (see Discussion) that gene flow is an active evolutionary force that prevents the genetic isolation of wild and domesticated population.

Analysis of genetic differentiation and gene flow among populations of Chiapas

To obtain an indication of the level of gene flow between wild and domesticated populations, we studied the Chiapas sample in more detail because information on the different degrees of isolation between wild and domesticated landrace populations was available for this sample from the exploration conducted as part of this research (see Materials and methods). Twenty populations originated in three different geographical areas, Tuxtla, Teopisca, and Las Rosas (Fig. 1). Sympatric wild and domesticated populations occurred in two of the three regions, with one pair in Tuxtla (site A) and three in Teopisca (sites E_E, E_S, and I). Parapatric populations were distributed in the Teopisca region (sites E, I, and J) and the Las Rosas region (site M).

Genetic diversity among and within populations was assessed with 145 AFLP markers of which 101 were polymorphic (Papa and Gepts 2003). After excluding loci putatively under the effect of selection (see Material and methods), an assessment of the level of genetic differentiation was obtained using 78 markers from pairwise estimates of F_{ST} between wild and domesticated populations at different levels of spatial proximity. F_{ST} values (Table 6) were the lowest for pairwise comparisons among domesticated populations (0.30), followed by those among wild populations (0.48) and those between wild and domesticated populations (0.60), regardless of their spatial proximity (sympatry, parapatry, or allopatry) (Table 6). The average pairwise estimate of F_{ST} between wild and domesticated populations in sympatry was 0.44. According to a Wilcoxon nonparametric test (Sokal and Rohlf 1995), it was significantly lower than the average F_{ST} between populations in parapatry (0.56; $P = 0.016$) and allopatry (0.63; $P = 0.003$) (Table 6). Differences in F_{ST} values between wild and domesticated populations in parapatry and allopatry were also significant ($P < 0.05$). The apparent lack of isolation in sympatry is also in agreement with the significantly higher estimates—as determined by a Wilcoxon nonparametric test ($P < 0.05$)—of within-population diversity observed for the populations in close-range sympatry (wild: Hs = 0.128; domesticated: Hs = 0.103) compared to

Table 5 Non-parametric (Kendall's) correlation between linkage distances of markers (cM) to the nearest domestication locus and F_{ST} or genetic diversity (H_{Wild}, H_{Dom}, and H_{Total})

Sample	No. of markers		F_{ST}	H_{Wild}	H_{Dom}	H_{Total}
Pooled	32	tau	−0.395	−0.268	0.032	−0.402
		P	0.002	0.037	NS	0.002
Chiapas	15	tau	−0.527	−0.522	−0.138	−0.539
		P	0.007	0.007	NS	0.006
Mesoamerica	17	tau	−0.181	−0.053	0.133	−0.190
		P	NS	NS	NS	NS

Considering linkage groups B01, B02, B03, B04, B06, B07, B08 and B11

Table 6 Average F_{ST} values for pairwise comparisons between individual domesticated and wild populations of common bean (*P. vulgaris*)

| | Pairwise F_{ST} | |
	Number of comparisons	Average[a]
Overall	96	0.60
Allopatry	70	0.63
Parapatry	22	0.56
Sympatry	4	0.44
Average	Not applicable	0.54

[a]Differences among average F_{ST} values for each spatial comparison were significant according to a Wilcoxon non-parametric test (see text)

values for the allopatric and parapatric populations (wild: Hs = 0.082; domesticated: Hs = 0.069).

Discussion

Many reports have been published that describe differences for F_{ST} at different loci. Increasingly, these differences can be correlated with specific genes, candidate genes, or QTLs subject to selection (reviewed in Storz 2005). Based on its predominantly self-pollinated reproductive biology and history of domestication, we anticipated the following genomic distribution of F_{ST} and H values. We expected that wild and domesticated populations had remained isolated from each other to a large extent from each other since domestication, which took place at least 2,500 years ago (Kaplan and Lynch 1999). The divergence between the two populations would have been mainly due to drift and selection during and after domestication in the domesticated gene pool. We expected differences for F_{ST} across the genome, with values in genome regions around the domestication genes (D markers) larger than those in genome regions not directly selected during domestication (UN and ND markers). The size of the genomic regions subject to hitchhiking around genes for domestication would be proportional to the ratio of selection intensity over the effective recombination frequency. With regard to genetic diversity, we expected non-significant differences for H across the entire genome (UN, ND, and D markers) in the wild population. In the domesticated populations, we expected a statistically significant reduction in genetic diversity for all the marker categories (due to the domestication bottleneck) and, potentially, an additional reduction in diversity for the D markers (selection during domestication), depending on the resolution of our analysis and the level of effective recombination in the domesticated gene pool. The genetic bottleneck induced by domestication has been well-documented in common bean (Gepts et al. 1986; Sonnante et al. 1994), as have the map locations of genes controlling the traits responsible for the domestication syndrome (Koinange et al. 1996).

Our results fit this expectation for F_{ST} values but not for H values. To explain this discrepancy with our initial assumption, we now offer an alternative explanation, namely that introgression, predominantly from domesticated wild types (Papa and Gepts 2003), provides a scenario consistent with the F_{ST} and H values presented. Recent research in Mexico has shown that gene flow can take place between wild and domesticated beans (González et al. 2005; Payró de la Cruz et al. 2005), with a three- to four-fold higher level from domesticated to wild types compared to the other direction (Papa and Gepts 2003; Zizumbo-Villarreal et al. 2005). The existence of gene flow between domesticated and wild types was further confirmed by our current data obtained in the Chiapas sample. Sympatric wild and domesticated populations displayed both lower differentiation and higher within-population diversity compared to parapatric and allopatric populations.

Asymmetric introgression, i.e., predominantly from the domesticated to the wild population (as determined by Papa and Gepts 2003), could explain why the levels of H for UN and ND markers in wild populations were low and not significantly different from those in the domesticated populations except in the Chiapas sample for ND markers (Fig. 3). Such asymmetric gene flow could be due to the much larger pollen load of the domesticated, compared to the wild population. In addition, it could be due to differential selection against hybrids, and the timing thereof, in wild versus domesticated populations. Wild traits are usually dominant or partially dominant in common bean (Koinange et al. 1996). Hence, F_1 and later generation hybrids will be phenotypically more similar to their wild than to their domesticated progenitors. This situation may lead to differential selection pressures in the two contrasting cultivated and wild environments (Papa and Gepts 2004). For example, farmers may exercise a strong selection pressure against F_1 domesticated × wild hybrids, for example against certain seed colors, shapes, and sizes, because of their lack of consumer appeal (Zizumbo-Villarreal et al. 2005). This situation may limit the entry of wild gametes in domesticated populations, in effect acting like a postzygotic reproductive isolation barrier. In wild populations, in contrast, F_1 plants resulting from wild × domesticated hybridization should have a higher fitness because of the milder selection (due to the predominantly wild phenotype of hybrids). Therefore, domesticated alleles could be more likely transmitted to the F_2 and later generations in wild populations; recombination in the hybrid progenies would dissipate the link between markers and domestication genes and, therefore, lead to the observed negative correlation between linkage distance and H or F_{ST} in wild but not domesticated populations as we observed (Table 5). This situation suggests that selection against alleles of the domestication syndrome is acting mainly in the wild environment in the segregating generations after hybridization.
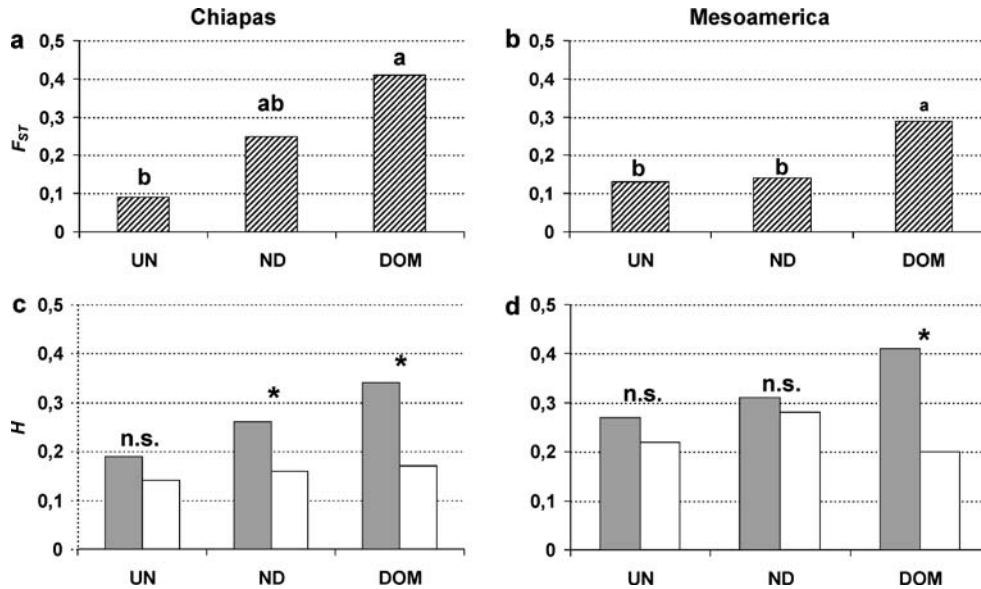
Fig. 3 Average population differentiation (F_{ST}) and unbiased gene diversity (H) in *P. vulgaris*. **a, b** F_{ST} between wild and domesticated populations of Chiapas and Mesoamerica, respectively; **c, d** H in wild (*clear bars*) and domesticated (*shaded bars*) populations of Chiapas and Mesoamerica, respectively. *DOM and ND* markers linked to genes for domestication and other traits, respectively; *UN* markers unlinked to known genes. *Asterisk and different letters* indicate significant differences ($P < 0.05$) with the Wilcoxon nonparametric test. For F_{ST}, significances were obtained after Bonferroni correction

The repeated gene flow over the years of sympatry from genetically depauperate domesticated populations to initially more diverse wild populations could lead to a displacement of the native genetic diversity in wild populations. This genetic assimilation could also have affected regions around domestication genes, if it were not for the fact that selection could act against these maladaptive domesticated alleles introduced by gene flow. Under a migration-selection balance, genetic assimilation occurs when the migration rate exceeds the selection coefficient (Lenormand 2002). In areas of the genome unlinked to D loci, no selection against domesticated alleles will occur and, if preferential introgression from domesticated into wild types is assumed, even small rates of migration will exceed the selection coefficient. At neutral loci linked to D loci, introgression will be reduced because of the parallel reduction of effective recombination imposed by linkage. Thus, the rate of introgression will be gradually reduced closer to D loci until a threshold is reached where the rate of introgression will equal the selection coefficient. The fact that the highest F_{ST} values were observed around domestication loci, therefore, reflects both positive selection for domestication alleles in the domesticated gene pool and background selection against maladaptive domesticated alleles introduced in the wild populations. As a consequence of the process described here, genetic assimilation will take place when alleles in wild populations are replaced by alleles from domesticated populations. Because domesticated populations are less variable than wild populations, genetic assimilation will also lead to a reduction in genetic diversity in wild populations. Therefore, gene flow from domesticated populations presents a threat to the continued existence of genetic diversity in wild populations (except in the genome regions around domestication genes).

Results of the two samples—Chiapas and Mesoamerica—differed, however, in the level of differentiation around non-domestication genes, principally disease resistance genes as detected by ND markers. In the geographically narrower Chiapas sample, differentiation measured by the ND markers was intermediate between that for the D and UN markers (Fig. 3a). In the geographically more widespread Mesoamerica sample, in contrast, differentiation detected by the ND and UN markers were not significantly different from each other but were significantly lower than that revealed by D markers (Fig. 3b). A potential explanation for this discrepancy may lie in the respective scope of the two samples as discussed recently by Lin et al. (2002). In the Chiapas sample, hitchhiking effects between genes under selection for particular traits important in local adaptation such as disease resistance and markers may have played a role in the observed differentiation between wild and domesticated types. In the Mesoamerica sample, local selection may also have occurred but its effects may have averaged out over the entire geographic area covered by the sample. Alternatively, the more ancient nature of the Mesoamerican sample would have led to more opportunities for recombination and reduced hitchhiking compared to the Chiapas sample.

Our results provide proof-of-concept support for the conduct of genome scans or signature-of-selection mapping to locate genes involved in adaptation on the molecular linkage map of a species (Schlötterer 2003;

Storz 2005). One of the advantages of genome scans is that they can confirm—at the population level—the map location of genes for previously identified traits. For example, Koinange et al. (1996) identified a QTL with possibly pleiotropic effects on growth habit and phenology on linkage group B08 of the map developed in a cross between a domesticated and a wild bean (Fig. 2). To date, no specific, discrete morphological or physiological trait corresponding to this QTL has been identified. However, the current analysis, based on F_{ST} and H values, confirms the importance of such a gene as a distinguishing factor between wild and domesticated beans among two broader samples of bean genotypes (Table 2). Although formal measures of LD distances have not been made as yet in common bean, distances of several tens or thousands of base pairs have been observed in other selfing species (Nordborg et al. 2002). Distances of this order of magnitude preclude precise mapping by LD. Hence, in a predominantly selfing species like common bean, genome scans or hitchhiking mapping may be a precursor to linkage mapping in segregating populations of known pedigree. This is in contrast with outcrossing species where LD mapping may represent the final step towards identification of the causal DNA polymorphism for a specific trait.

In the end, gene flow between the common bean and its wild progenitor will take place (Papa and Gepts 2003; González et al. 2005; Payró de la Cruz et al. 2005; Zizumbo-Villarreal et al. 2005; and current results). Wild and domesticated beans have been sympatric for at least 2,500 years (Kaplan and Lynch 1999). During this period, divergent selection appears to have been a major evolutionary factor maintaining the identity of sympatric wild and domesticated populations in the face of gene flow between them. Further research is necessary to determine the year-to-year and location-to-location dynamics of common bean in its center of origin. Our current findings provide an experimental framework to assess the long-term risk of transgene escape. Given enough time, transgenes will be transferred by gene flow into populations of wild relatives even in predominantly selfing species. However, the probability of survival of such transgenes depends on their location in the genome, in addition to other factors such as their selective value. Our results provide some experimental support for a transgene mitigation strategy whereby transgenes located at or tightly linked to loci, subject to disruptive selection, will be indirectly selected against, thus decreasing their chances of becoming established in populations of wild crop relatives (Gressel 1999).

References

Burke JM, Tang S, Knapp SJ, Rieseberg LH (2002) Genetic analysis of sunflower domestication. Genetics 161:1257–1267

Cavalli-Sforza LL (1966) Population structure and human evolution. Proc R Soc Lond Ser B 164:362–379

Charlesworth B, Nordborg M, Charlesworth D (1997) The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. Genet Res 70:155–174

Doebley J, Stec A (1991) Genetic analysis of the morphological differences between maize and teosinte. Genetics 129:285–295

Doebley J, Stec A, Wendel J, Edwards M (1990) Genetic and morphological analysis of a maize-teosinte F_2 population:implications for the origin of maize. Proc Natl Acad Sci USA 87:9888–9892

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

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 map and alignment of RFLP maps. Theor Appl Genet 97:847–856

Geffroy V, Sévignac M, De Oliveira J, Fouilloux G, Skroch P, Thoquet P, Gepts P, Langin T, Dron M (2000) Inheritance of partial resistance against *Colletotrichum lindemuthianum* in *Phaseolus vulgaris* and co-localization of QTL with genes involved in specific resistance. Mol Plant Microb Interact 13:287–296

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

Gepts P (1999) Development of an integrated genetic linkage map in common bean (*Phaseolus vulgaris* L.) and its use. In: Singh S (ed) Bean breeding for the 21st century Kluwer, Dordrecht, pp 53–91, 389–400

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

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

González A, Wong A, Delgado-Salinas A, Papa R, Gepts P (2005) Assessment of inter simple sequence repeat markers to differentiate sympatric wild and domesticated populations of common bean (*Phaseolus vulgaris* L.). Crop Sci 35:606–615

Gressel J (1999) Tandem constructs: preventing the rise of superweeds. Trends Biotech 17:361–366

Ibarra-Pérez F, Ehdaie B, Waines G (1997) Estimation of outcrossing rate in common bean. Crop Sci 37:60–65

Kaplan L, Lynch T (1999) *Phaseolus* (Fabaceae) in archaeology: AMS radiocarbon dates and their significance for pre-Columbian agriculture. Econ Bot 53:261–272

Koenig R, Gepts P (1989) Allozyme diversity in wild *Phaseolus vulgaris*: further evidence for two major centers of 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

Lenormand T (2002) Gene flow and the limits to natural selection. Trends Ecol Evol 17:183–189

Lin J-Z, Morrell PL, Clegg MT (2002) The influence of linkage and inbreeding on patterns of nucleotide sequence diversity at duplicate alcohol dehydrogenase loci in wild barley (*Hordeum vulgare* ssp. *spontaneum*). Genetics 162:2007–2015

Llaca V, Gepts P (1996) Pulsed field gel electrophoresis analysis of the phaseolin locus region in *Phaseolus vulgaris*. Genome 39:722–729

McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. Trends Ecol Evol 17:285–291

Merilä J, Crnokrak P (2001) Comparison of genetic differentiation at marker loci and quantitative traits. J Evol Biol 14:892–903

Miklas P, Johnson W, Delorme R, Gepts P (2001) QTL conditioning physiological resistance and avoidance to white mold in dry bean. Crop Sci 41:309–315

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York, p 512

Nei M, Li W-H (1979) Mathematical models for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Nodari RO, Tsai SM, Guzmán P, Gilbertson RL, Gepts P (1993) Towards an integrated linkage map of common bean. 3. Mapping genetic factors controlling host-bacteria interactions. Genetics 134:341–350

Nordborg M, Borevitz JO, Bergelson J, Berry CC, Chory J, Hagenblad J, Kreitman M, Maloof JN, Noyes T, Oefner PJ, Stahl EA, Weigel D (2002) The extent of linkage disequilibrium in Arabidopsis thaliana. Nat Genet 30:190–193

Papa R, Gepts P (2003) Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (Phaseolus vulgaris L.) from Mesoamerica. Theor Appl Genet 106:239–250

Papa R, Gepts P (2004) Asymmetric gene flow and introgression between wild and domesticated populations. In: Den Nijs D, Bartsch D, Sweet J (ed) Introgression from genetically modified plants into wild relatives and its consequences. CABI, Oxon, pp 125–138

Payró de la Cruz E, Gepts P, Colunga GarciaMarín P, Zizumbo Villareal D (2005) Spatial distribution of genetic diversity in wild populations of Phaseolus vulgaris L. from Guanajuato and Michoacán, México. Genet Res Crop Evol (in press)

Poncet V, Lamy F, Enjalbert J, Joly H, Sarr A, Robert T (1998) Genetic analysis of the domestication syndrome in pearl millet (Pennisetum glaucum L, Poaceae): inheritance of the major characters. Heredity 81:648–658

Poncet V, Lamy F, Devos K, Gale M, Sarr A, Robert T (2000) Genetic control of domestication traits in pearl millet (Pennisetum glaucum L., Poaceae). Theor Appl Genet 100:147–159

Poncet V, Martel E, Allouis S, Devos K, Lamy F, Sarr A, Robert T (2002) Comparative analysis of QTLs affecting domestication traits between two domesticated × wild pearl millet (Pennisetum glaucum L., Poaceae) crosses. Theor Appl Genet 104:965–975

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

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

Sokal R, Rohlf F (1995) Biometry: the principles and practice of statistics in biological research, 3rd edn. Freeman, New York

Sonnante G, Stockton T, Nodari RO, Becerra Velásquez VL, Gepts P (1994) Evolution of genetic diversity during the domestication of common-bean (Phaseolus vulgaris L.). Theor Appl Genet 89:629–635

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

Tenaillon MI, U'Ren J, Tenaillon O, Gaut BS (2004) Selection versus demography: a multilocus investigation of the domestication process in maize. Mol Biol Evol 21:1214–1225

Vigouroux Y, McMullen 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

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

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

Xiong L, Liu K, Dai X, Xu C, Zhang Q (1999) Identification of genetic factors controlling domestication-related traits of rice using an F_2 population of a cross between Oryza sativa and O. rufipogon. Theor Appl Genet 98:243–251

Zizumbo-Villarreal D, Colunga-GarcíaMarín P, Payró de la Cruz E, Delgado-Valerio P, Gepts P (2005) Population structure and evolutionary dynamics of wild–weedy–domesticated complexes of common bean in a Mesoamerican region. Crop Sci 35 (in press)