

The genetic make-up of the European landraces of the common bean

S. A. Angioi¹, D. Rau¹, L. Nanni², E. Bellucci², R. Papa^{2,3} and G. Attene^{1,4*}

¹Dipartimento di Scienze Agronomiche e Genetica Vegetale Agraria, Università degli Studi di Sassari, Sassari, Italy, ²Dipartimento di Scienze Ambientali e delle Produzioni Vegetali, Università Politecnica delle Marche, Ancona, Italy, ³CRA-CER Council for Agricultural Research – Cereal Research Centre, S.S. 16, Km 675, 71122 Foggia, Italy and ⁴Centro per la Conservazione e Valorizzazione della Biodiversità Vegetale, Università degli Studi di Sassari, Surigheddu, Alghero, Italy

Abstract

Here, we present a brief overview of the main studies conducted on the common bean (*Phaseolus vulgaris* L.) in Europe and other countries outside its centres of origin. We focus on the proportions of the Andean and Mesoamerican gene pools, and on the inter-gene pool hybridization events. In Europe, for chloroplast microsatellites, 67% of European germplasm is of Andean origin. Within Europe, interesting trends have been seen; indeed, the majority of the Andean type is found in the three macro-areas of the Iberian Peninsula, Italy and central-northern Europe, while, in eastern and south-eastern Europe, the proportion of the Mesoamerican type increased. On a local scale, the contribution of the Mesoamerican type is always low. On other continents, various situations are seen using different markers: in China and Brazil, the Mesoamerican gene pool prevails, while in an African sample, overall, both gene pools are equally represented, with differences in individual countries. The frequency of European bean genotypes deriving from at least one hybridization event was 44% with an uneven distribution. Interestingly, hybrids tend to have intermediate seed size in comparison with ‘pure’ Andean or Mesoamerican types. On other continents, very few hybrids are found, probably because of the different marker systems used.

Keywords: Introgression; Mesoamerican and Andean gene pools; *Phaseolus vulgaris*

Introduction

The common bean (*Phaseolus vulgaris* L.) is of great agronomic interest worldwide, and represents 50% of grain legumes for direct human consumption (McClellan *et al.*, 2004). Domestication of *P. vulgaris* occurred independently in the Mesoamerican and Andean areas, which gave rise to two highly differentiated gene pools (Gepts and Debouck, 1991; Gepts, 1998). The two gene pools can be distinguished according to morphological traits (Gepts *et al.*, 1986), phaseolins (major

seed-storage proteins) and by various marker systems (Beebe *et al.*, 2001; Papa and Gepts, 2003; Kwak and Gepts, 2009; Rossi *et al.*, 2009). The Mesoamerican types are small or medium seeded, with phaseolins S or B; while the Andean are large seeded, with phaseolins T, C, H and A (Gepts *et al.*, 1986; Singh *et al.*, 1991). Based on the occurrence of a strong bottleneck in the Andean wild population, Rossi *et al.* (2009) recently suggested a Mesoamerican origin of the common bean.

The common bean was introduced into Europe after Columbus’s voyages. It was distributed widely in all parts of Europe, where many landraces and varieties evolved, as they were grown to provide dry seeds or fresh pods (Zeven, 1997).

*Corresponding author. E-mail: attene@uniss.it

Here, we present a brief overview regarding the level and structure of the genetic diversity of the European common bean, and we compare this with information provided for other continents.

Composition of the European common bean

Overall, outside the domestication centres of the common bean (Table 1), the proportions of the Andean and Mesoamerican gene pools vary considerably across different countries and continents.

In Europe, studies have been conducted on different scales: continental, macro-areas, single country and local (within country). The first studies on large collections of the European common bean were carried out by using phaseolins. These demonstrated that the germplasm arose from both of the American gene pools, with a higher frequency of Andean types (76–66%) (Gepts and Bliss, 1988; Lioi, 1989). This prevalence (76%) was confirmed in a large collection that included, for the first time, central European countries (Logozzo *et al.*, 2007).

Angioi *et al.* (2010) used chloroplast microsatellites (cpSSRs), two nuclear markers and morphological seed traits to analyze a large part of the collection of Logozzo *et al.* (2007), while adding new accessions from countries previously less represented (e.g. France). Also for this analysis, the prevalence (67%) of European germplasm was of Andean origin (Fig. 1).

Within Europe, an interesting trend was seen, as the Andean type was in the majority in three macro-areas: the Iberian Peninsula, Italy and central-northern Europe. In contrast, in eastern and south-eastern Europe, the proportion of the Mesoamerican type increased (Fig. 1). This was supported by other studies in the Iberian Peninsula (Rodrigo *et al.*, 2003; Ocampo *et al.*, 2005), and at a country level in Greece (46%; Lioi, 1989) and Bulgaria (79%; Svetleva *et al.*, 2006) (see Papa *et al.*, 2006, for review).

In Italy, on a local scale, the prevalence of the Andean type has been confirmed. The contribution of the Mesoamerican gene pool varied from 5% in Sardinian landraces (Angioi *et al.*, 2009) to 29% in the Marche region (Sicard *et al.*, 2005), according to chloroplast and nuclear markers, and 12 and 13% in Abruzzo and Basilicata, respectively (Piergiorganni *et al.*, 2000a, b), according to phaseolins.

Looking at other continents, the Mesoamerican gene pool prevailed in a Chinese collection (Zhang *et al.*, 2008) and a Brazilian collection (Burle *et al.*, 2010) (Table 1 and Fig. 1). The Mesoamerican predominance in Brazil is surprising, given its close proximity to the Andes. Multiple introductions of Mesoamerican

Table 1. Gene pool compositions (% of Andean and Mesoamerican), gene diversities (Nei, 1978) and % hybrids in European sample, compared with countries on other continents

References	Sample size	Molecular/ biochemical marker	Gene pool composition (%)		Genetic diversity (<i>H</i>)		
			Andean	Mesoamerican	Andean	Mesoamerican	Hybrids (%)
Europe							
Angioi <i>et al.</i> (2010)	307	cpSSR ^a	67	33	0.36	0.34	44 ^b
Brazil							
Burle <i>et al.</i> (2010)	279	nuSSR ^c	21	79	0.30	0.33	4 ^d
China							
Zhang <i>et al.</i> (2008)	299	nuSSR	25	75	0.42	0.37	5 ^e
Ethiopia							
Asfaw <i>et al.</i> (2009)	99	nuSSR	28	72	0.44	0.59	1 ^f
Kenya							
Asfaw <i>et al.</i> (2009)	89	nuSSR	73	27	0.45	0.59	0
Rwanda							
Blair <i>et al.</i> (2010)	355	nuSSR	40	60	–	–	10 ^f
East Africa ^g							
Cepts and Bliss (1988)	111	phaseolin	81	19	–	–	–

^aProvan *et al.* (2001), for a review. ^bData obtained through comparison with nuclear STS and phaseolins. ^cAs defined by Tautz 1998. ^dData obtained through comparison with phaseolins. ^eData obtained through comparison with morphological data. ^fIntermediate genotypes identified in the neighbour-joining tree. ^gEthiopia, Kenya, Uganda, Burundi, Rwanda, Democratic Republic of Congo (former Zaire), Tanzania, Malawi and Zambia.

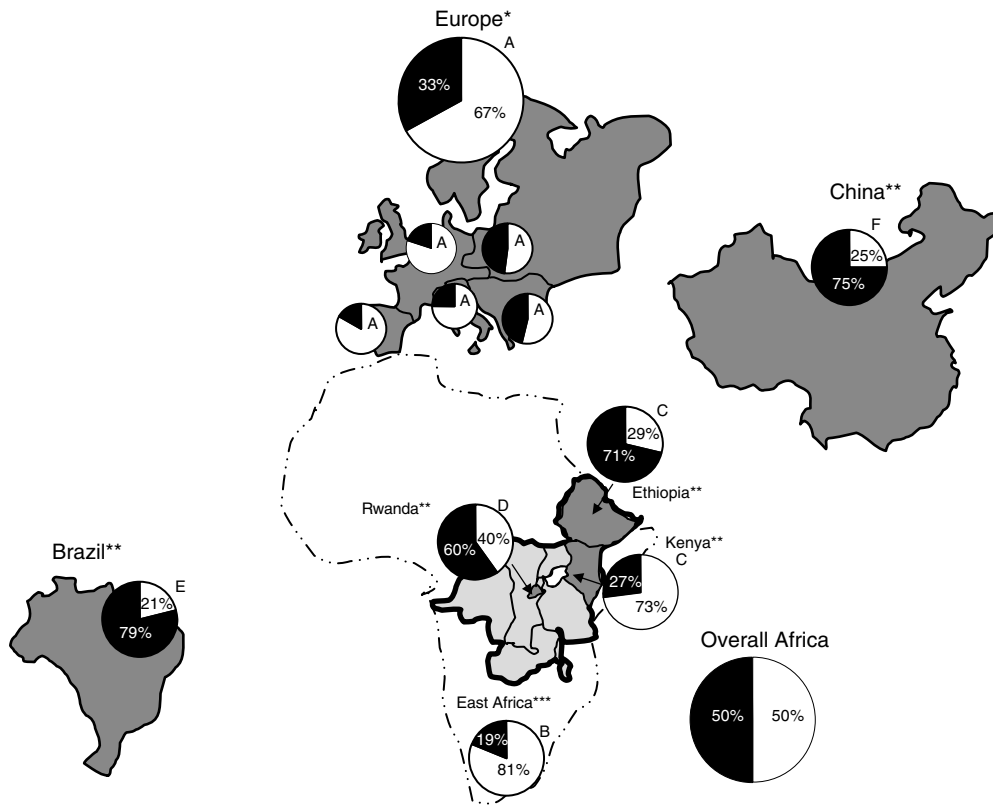


Fig. 1. Distribution map of the Andean and Mesoamerican gene pools in Europe and on other continents, analyzed with molecular markers. In the pie charts: white, Andean gene pool; black, Mesoamerican gene pool. (A) Europe (sample size, $n = 307$) and the Iberian Peninsula (53), Italy (32), central-northern Europe (74), eastern Europe (69), south-eastern Europe (79); Angioi *et al.* (2010). (B) east Africa (111), Gepts and Bliss (1988); (C) Ethiopia (99) and Kenya (89), Asfaw *et al.* (2009); (D) Rwanda (355), Blair *et al.* (2010); (E) Brazil (279), Burle *et al.* (2010); (F) China (299), Zhang *et al.* (2008). *cpSSRs, **nuSSRs, ***phaseolins. The overall African pie chart is obtained pooling together the data provided in Table 1.

germplasm and similarities in climate and soil between Brazil and Mesoamerica might have had a considerable impact in establishing this pattern (Burle *et al.*, 2010).

In Africa, overall, both gene pools are equally represented (Fig. 1), although at the single country level, contrasting situations are seen. In Ethiopia (Asfaw *et al.*, 2009) and Rwanda (Blair *et al.*, 2010), the Mesoamerican types predominate, and *vice versa* in Kenya (Asfaw *et al.*, 2009) and East Africa (Gepts and Bliss, 1988) (Table 1 and Fig. 1). This suggested that the genetic divergence in bean landraces might be due to the original differences in the introduced germplasm from the centres of origin (Asfaw *et al.*, 2009), combined with differences in pest resistance (e.g. Mesoamerican types resistant to root rot in Rwanda) and production (Mesoamerican genotypes have the highest yields).

Finally, gene flow between Ethiopia and Kenya was moderate, probably due to different farmer preferences according to ecological adaptation, cooking values and market orientation (Asfaw *et al.*, 2009). This is in contrast to Europe. Indeed, in Europe, high gene flow among

macro-areas and/or homogenizing selection (anthropic and natural) has been suggested (Angioi *et al.*, 2010). The role of selection is suggested by the findings of Logozzo *et al.* (2007) in European landraces, where accessions with Mesoamerican phaseolin had significantly larger seed size than individuals from America in the same phaseolin class.

Introgression between the gene pools

Introgression is an event that arises from hybridization among gene pools, as spontaneous outcrossing in farmer fields followed by selection for adaptation to production niches and uses. Comparing chloroplast data with nuclear (phaseolin and sequence-tagged sites, STS) and morphological data, Angioi *et al.* (2010) estimated that a high proportion of the European bean germplasm (44%) derived from at least one hybridization event using a maximum-likelihood approach. Although hybrids are present everywhere, they show uneven distributions,

with high frequencies in central Europe and low frequencies in the Iberian Peninsula and Italy. A comparison of chloroplast data with nuclear and morphological data is a reliable method to identify the hybrids, as they tend to have intermediate seed size with respect to 'pure' Andean or Mesoamerican, with Andean × Mesoamerican seeds smaller than 'pure' Andean, and Mesoamerican × Andean seeds larger than 'pure' Mesoamerican. This method was also applied at local scales in Italy, in the Marche region (12% hybrids; Sicard *et al.*, 2005) and Sardinia (4%; Angioi *et al.*, 2009).

Other studies have analyzed hybridization among gene pools, but they found few hybrids: 4% in Brazil, comparing nuclear SSRs (nuSSRs) to phaseolins (Burle *et al.*, 2010), and from 1 to 10% in Ethiopia and Rwanda (Asfaw *et al.*, 2009; Blair *et al.*, 2010), considering individuals intermediate among gene pools in the neighbour-joining tree (Table 1). The differences in hybrid frequency can be explained by different marker systems used (chloroplast and nuclear) and the definition of hybrids as recent (Asfaw *et al.*, 2009; Blair *et al.*, 2010) versus old generation hybrids (Angioi *et al.*, 2010). Another explanation might be that, as seen in Brazil, the frequency of the two gene pools are very different, or that in some environments, there is no flowering synchronization between Andean and Mesoamerican types (Asfaw *et al.*, 2009). In the Chinese sample, Zhang *et al.* (2008) found 5% hybrids, noting that average seed weight of the Andean types was lower than that of the American Andean beans, with the opposite for the Mesoamerican Chinese bean.

The existence of the high frequency of inter-gene pool hybridization in Europe might have had a significant impact on the structure of the genetic and genotypic diversity in the nuclear genome. This is consistent with the breakdown of geographical isolation between the two gene pools (Angioi *et al.*, 2010). Moreover, the European hybrids appear to be of great importance for breeding that aims to recombine Andean and Mesoamerican traits (Johnson and Gepts, 1999, 2002).

Acknowledgements

This study was partially supported by the Italian Government (MIUR) grant no. 20083PFSXA PRIN 2008.

References

- Angioi SA, Desiderio F, Rau D, Bitocchi E, Attene G and Papa R (2009) Development and use of chloroplast microsatellites in *Phaseolus* spp. and other legumes. *Plant Biology* 11: 598–612.
- Angioi SA, Rau D, Attene G, Nanni L, Bellucci E, Logozzo G, Negri V, Spagnoletti Zeuli PL and Papa R (2010) Beans in Europe: origin and structure of the European landraces of *Phaseolus vulgaris* L. *Theoretical and Applied Genetics* 121: 829–843 doi: 10.1007/s00122-010-1353-2.
- Asfaw A, Blair MW and Almekinders C (2009) Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) landraces from the East African highlands. *Theoretical and Applied Genetics* 120: 1–12.
- Beebe S, Rengifo J, Gaitan E, Duque MC and Tohme J (2001) Diversity and origin of Andean landraces of common bean. *Crop Science* 41: 854–862.
- Blair MW, González LF, Kimani M and Butare L (2010) Genetic diversity, inter-gene pool introgression and nutritional quality of common beans (*Phaseolus vulgaris* L.) from Central Africa. *Theoretical and Applied Genetics* 121: 237–248.
- Burle ML, Fonseca JR, Kami JA and Gepts P (2010) Microsatellite diversity and genetic structure among common bean (*Phaseolus vulgaris* L.) landraces in Brazil, a secondary center of diversity. *Theoretical and Applied Genetics* 121: 801–813 doi: 10.1007/s00122-010-1350-5.
- Gepts P (1998) Origin and evolution of common bean: past events and recent trends. *Horticultural Science* 33: 1121–1130.
- Gepts P and Bliss FA (1988) Dissemination pathways of common bean (*Phaseolus vulgaris*, Fabaceae) deduced from phaseolin electrophoretic variability. II Europe and Africa. *Economic Botany* 42: 86–104.
- Gepts P and Debouck DG (1991) Origin, domestication and evolution of common bean, *Phaseolus vulgaris*. In: Schoonhoven A and Voysey O (eds) *Common Beans: Research for Crop Improvement*. CAB International Wallingford, UK, pp 4–54.
- Gepts P, Osborne TC, Rashka K and Bliss FA (1986) Electrophoretic analysis of phaseolin protein variability in wild forms and landraces of the common bean *Phaseolus vulgaris* L.: evidence for two centers of domestications. *Economic Botany* 40: 451–468.
- Johnson WC and Gepts P (1999) Segregation for performance in recombinant inbred populations resulting from inter-gene pool crosses of common bean (*Phaseolus vulgaris* L.). *Euphytica* 106: 5–56.
- Johnson WC and Gepts P (2002) The role of epistasis in controlling seed yield and other agronomic traits in an Andean × Mesoamerican cross of common bean (*Phaseolus vulgaris* L.). *Euphytica* 125: 69–79.
- Kwak M and Gepts P (2009) Structure of genetic diversity in the two major gene pools of common bean (*Phaseolus vulgaris* L., Fabaceae). *Theoretical and Applied Genetics* 118: 979–992.
- Lioi L (1989) Geographical variation of phaseolin patterns in an old world collection of *Phaseolus vulgaris*. *Seed Science and Technology* 17: 317–324.
- Logozzo G, Donnoli R, Macaluso L, Papa R, Knupffer H and Spagnoletti Zeuli PL (2007) Analysis of the contribution of Mesoamerican and Andean gene pools to European common bean (*Phaseolus vulgaris* L.) germplasm and strategies to establish a core collection. *Genetic Resources and Crop Evolution* 54: 1763–1779.
- McClellan PE, Lee RK and Miklas PN (2004) Sequence diversity analysis of dihydroflavonol 4-reductase intron 1 in common bean. *Genome* 47: 266–280.
- Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583–590.

- Papa R and 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. *Theoretical and Applied Genetics* 106: 239–250.
- Papa R, Nanni L, Sicard D, Rau D and Attene G (2006) The evolution of genetic diversity in *Phaseolus vulgaris* L. In: Motley TJ, Zerega N and Cross H (eds) *New Approaches to the Origins, Evolution and Conservation of Crops. Darwin's Harvest*. New York: Columbia University Press.
- Piergiovanni AR, Cerbino D and Brandi M (2000a) The common bean populations from Basilicata (southern Italy). An evaluation of their variation. *Genetic Resources and Crop Evolution* 47: 489–495.
- Piergiovanni AR, Taranto G and Pignone D (2000b) Diversity among common bean populations from the Abruzzo region (central Italy): a preliminary inquiry. *Genetic Resources and Crop Evolution* 47: 467–470.
- Provan J, Powell W and Hollingsworth PM (2001) Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology and Evolution* 16: 142–147.
- Rodiño AP, Santalla M, De Ron AM and Singh SP (2003) A core collection of common bean from the Iberian peninsula. *Euphytica* 131: 165–175.
- Rossi M, Bitocchi E, Bellucci E, Nanni L, Rau D, Attene G and Papa R (2009) Linkage disequilibrium and population structure in wild and domesticated populations of *Phaseolus vulgaris* L. *Evolutionary Applications* 2: 504–522.
- Sicard D, Nanni L, Porfiri O, Bulfon D and Papa R (2005) Genetic diversity of *Phaseolus vulgaris* L and *P. coccineus* L. landraces in central Italy. *Plant Breeding* 124: 464–472.
- Singh SP, Nodari R and Gepts P (1991) Genetic diversity in cultivated common bean. I. Allozymes. *Crop Science* 31: 19–23.
- Svetleva D, Pereira G, Carlier J, Cabrita L, Leitão J and Genchev D (2006) Molecular characterization of *Phaseolus vulgaris* L. genotypes included in Bulgarian collection by ISSR and AFLP™ analyses. *Scientia Horticulturae* 109: 198–206.
- Tautz D (1998) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research* 17: 6463–6471.
- Zeven AC (1997) The introduction of the common bean (*Phaseolus vulgaris* L.) into Western Europe and the phenotypic variation of dry bean collected in the Netherlands in 1946. *Euphytica* 94: 319–328.
- Zhang X, Blair MW and Wang S (2008) Genetic diversity of Chinese common bean (*Phaseolus vulgaris* L.) landraces assessed with simple sequence repeats markers. *Theoretical and Applied Genetics* 117: 629–640.