



# Beans (*Phaseolus* ssp.) as a Model for Understanding Crop Evolution

Elena Bitocchi<sup>1</sup>, Domenico Rau<sup>2</sup>, Elisa Bellucci<sup>1</sup>, Monica Rodriguez<sup>2</sup>, Maria L. Murgia<sup>2</sup>, Tania Gioia<sup>3</sup>, Debora Santo<sup>1</sup>, Laura Nanni<sup>1</sup>, Giovanna Attene<sup>2</sup> and Roberto Papa<sup>1\*</sup>

<sup>1</sup> Department of Agricultural, Food and Environmental Sciences, Marche Polytechnic University, Ancona, Italy, <sup>2</sup> Department of Agriculture, University of Sassari, Sassari, Italy, <sup>3</sup> School of Agricultural, Forestry, Food and Environmental Sciences, University of Basilicata, Potenza, Italy

## OPEN ACCESS

### Edited by:

Nicolas Rispaill,  
Consejo Superior de Investigaciones  
Científicas (CSIC), Spain

### Reviewed by:

Zlatko Satovic,  
University of Zagreb, Croatia  
Klaudija Carović-Stanko,  
University of Zagreb, Croatia

### \*Correspondence:

Roberto Papa  
r.papa@univpm.it

### Specialty section:

This article was submitted to  
Crop Science and Horticulture,  
a section of the journal  
Frontiers in Plant Science

**Received:** 31 December 2016

**Accepted:** 19 April 2017

**Published:** 08 May 2017

### Citation:

Bitocchi E, Rau D, Bellucci E,  
Rodriguez M, Murgia ML, Gioia T,  
Santo D, Nanni L, Attene G and  
Papa R (2017) Beans (*Phaseolus*  
ssp.) as a Model for Understanding  
Crop Evolution.  
Front. Plant Sci. 8:722.  
doi: 10.3389/fpls.2017.00722

Here, we aim to provide a comprehensive and up-to-date overview of the most significant outcomes in the literature regarding the origin of *Phaseolus* genus, the geographical distribution of the wild species, the domestication process, and the wide spread out of the centers of origin. *Phaseolus* can be considered as a unique model for the study of crop evolution, and in particular, for an understanding of the convergent phenotypic evolution that occurred under domestication. The almost unique situation that characterizes the *Phaseolus* genus is that five of its ~70 species have been domesticated (i.e., *Phaseolus vulgaris*, *P. coccineus*, *P. dumosus*, *P. acutifolius*, and *P. lunatus*), and in addition, for *P. vulgaris* and *P. lunatus*, the wild forms are distributed in both Mesoamerica and South America, where at least two independent and isolated episodes of domestication occurred. Thus, at least seven independent domestication events occurred, which provides the possibility to unravel the genetic basis of the domestication process not only among species of the same genus, but also between gene pools within the same species. Along with this, other interesting features makes *Phaseolus* crops very useful in the study of evolution, including: (i) their recent divergence, and the high level of collinearity and synteny among their genomes; (ii) their different breeding systems and life history traits, from annual and autogamous, to perennial and allogamous; and (iii) their adaptation to different environments, not only in their centers of origin, but also out of the Americas, following their introduction and wide spread through different countries. In particular for *P. vulgaris* this resulted in the breaking of the spatial isolation of the Mesoamerican and Andean gene pools, which allowed spontaneous hybridization, thus increasing of the possibility of novel genotypes and phenotypes. This knowledge that is associated to the genetic resources that have been conserved *ex situ* and *in situ* represents a crucial tool in the hands of researchers, to preserve and evaluate this diversity, and at the same time, to identify the genetic basis of adaptation and to develop new improved varieties to tackle the challenges of climate change, and food security and sustainability.

**Keywords:** domestication, genetic diversity, adaptation, population genomics, crop evolution, convergent evolution, selection signatures

## INTRODUCTION

Beans (*Phaseolus* spp.), and in particular the common bean *P. vulgaris* L., represent the most important grain legume for direct human consumption worldwide. They are a major source of highly valuable plant protein and micronutrients (Broughton et al., 2003; Vaz Patto et al., 2015), they provide health benefits that are related to their regular consumption (Messina, 2014; Bitocchi et al., 2016b), and they contribute to sustainable improvements to the environment when they are grown in agricultural rotation or with intercropping, due to their biological nitrogen fixation, their effects on the soil, and their control of weeds (Rubiales and Mikic, 2015; Bitocchi et al., 2016b). Thus, beans have a key role in the diversification and sustainable intensification of agriculture, particularly in light of the new and urgent challenges, such as climate change. However, it has to be considered that without a deep knowledge of the evolutionary history of crops, no improvements are possible, inasmuch as evolutionary studies provide breeders with information about the available genetic diversity and the genetic control of important agronomic traits related to adaptation and domestication (Bitocchi et al., 2016b).

As Charles Darwin suggested, crop domestication can be seen as a “giant experiment” to test the evolutionary hypothesis. During domestication, similar sets of traits were selected over a wide range of plant species, as the so-called domestication syndrome, which shows numerous examples of convergent phenotypic evolution. *Phaseolus* ( $2n = 2X = 22$ ) is a unique example of multiple parallel and independent domestications. Indeed, not only did domestication occur in five closely related species, *Phaseolus vulgaris*, *P. lunatus*, *P. coccineus*, *P. dumosus* (formerly *P. polyanthus*), *P. acutifolius*, but also, different from other crop species, both *P. vulgaris* and *P. lunatus* have undergone two independent domestications. One was in Mesoamerica and the other in the Andes, which occurred during their reproductive isolation that was caused by the geographical barriers between these gene pools. Thus, considering single independent domestication events for *P. coccineus*, *P. dumosus* and *P. acutifolius*, and two for both *P. vulgaris* and *P. lunatus*, at least seven independent and isolated processes of domestication have occurred for *Phaseolus*.

A different question that arises is the occurrence of multiple domestications within species or within gene pools, where after domestication there was a lack of, or there were incomplete, reproductive barriers, with gene flow occurring among early domesticates, as can also be seen in for other crops (Meyer et al., 2012). In most of these cases where multiple domestications have occurred, this included gene flow between the early domesticates, while for beans, the strong geographical isolation between the gene pools guaranteed their reproductive isolation. However, this does not exclude *per se* the occurrence of multiple domestications within each gene pool of *P. lunatus* and *P. vulgaris*, and this topic is discussed later in this review. The only example that might be similar to *Phaseolus* was the domestication of rice, with the independent domestication of the *indica* and *japonica* subspecies in Asia (Vitte et al., 2004; Londo et al., 2006). However, recent studies appear to demonstrate a more complex situation for

rice, which involved a single domestication with the subsequent divergence of these two subspecies (Molina et al., 2011; Choi et al., 2017).

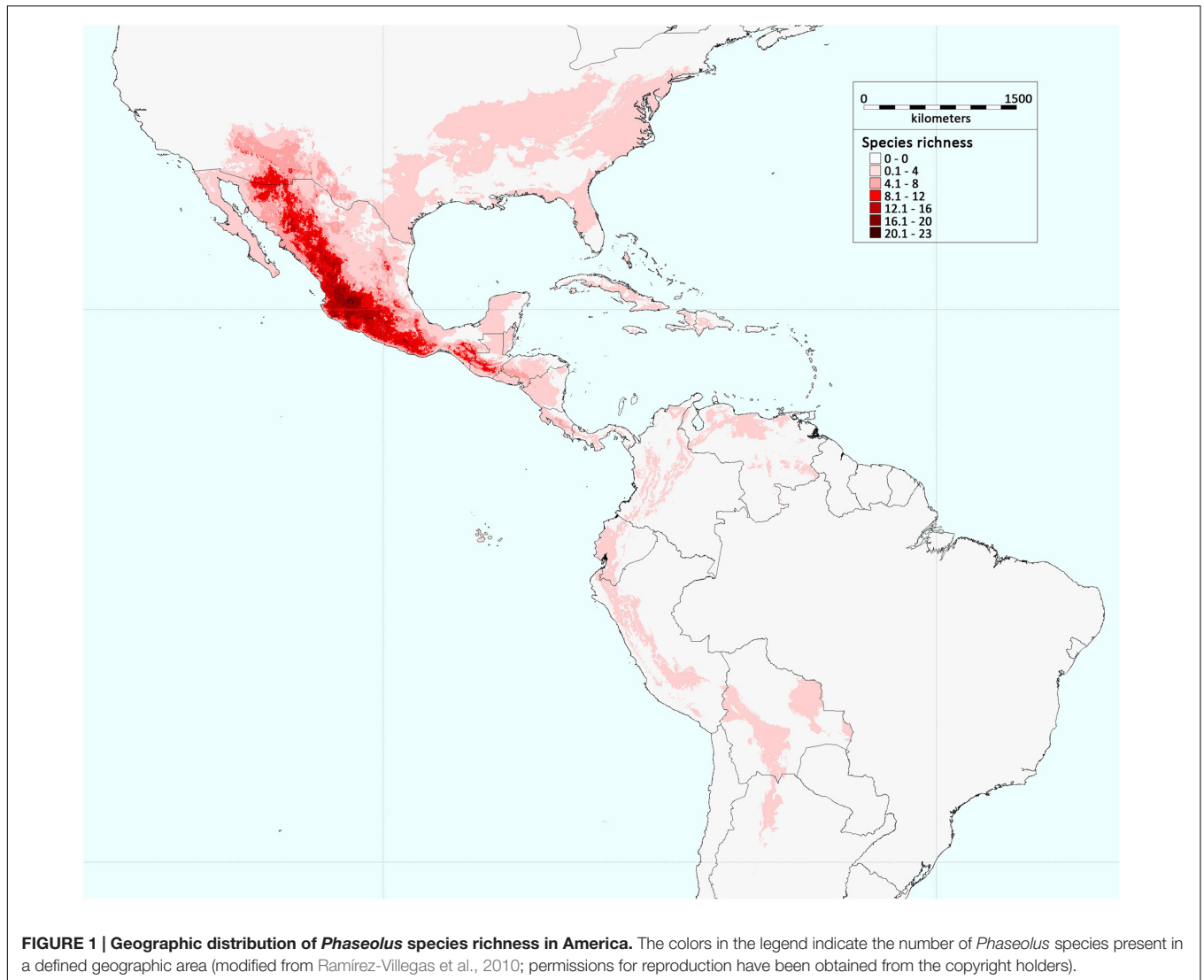
The *Phaseolus* genus also represents a very interesting model because the divergence among these five *Phaseolus* spp. is relatively recent (2–4 Ma ago; Delgado-Salinas et al., 2006). Indeed, genomic and cytogenetic analyses have reported high levels of collinearity and synteny among the *Phaseolus* genomes (Bonifácio et al., 2012; Fonsêca and Pedrosa-Harand, 2013; Gujaria-Verma et al., 2016), which suggests conserved gene function.

The *Phaseolus* spp. have different breeding systems and life history traits, from annual and autogamous, to perennial and allogamous. This provides the opportunity to determine whether these features have direct consequences on the effects of domestication on the phenotypic and genotypic architecture of the crop plants. Another important aspect, at least for the common bean, is the complex pattern of expansion and the pathways of distribution out of the American domestication centers. This also involved several introductions from the New World that were combined with exchanges between continents, and among several countries within continents. Thus, along with a dramatic increase in the amplitude of the agro-ecological conditions for these crops, new genetic combinations have allowed new opportunities for natural and human-mediated selection that might have promoted their adaptation to specific environmental conditions.

All of these features make *Phaseolus* spp. an ideal model to study domestication and evolution. Thus, the present review offers an overview of the current knowledge of the evolutionary history of the *Phaseolus* crop species, with particular focus on *P. vulgaris* L. and on the recent outcomes relating to the genetic bases of important domestication and adaptation traits.

## ORIGIN OF THE SPECIES OF THE *Phaseolus* GENUS

Among the ~70 species that belong to the *Phaseolus* genus, most are geographically distributed in Mesoamerica (Figure 1), where the genus appears to have diversified within the past 4–6 Ma ago (Delgado-Salinas et al., 2006). This diversification of the different species is likely to have taken place during and after the tectonic events that led to the present-day form of Mexico (Delgado-Salinas et al., 2006), which appeared in the Late Miocene (5 Ma ago; Nieto-Samaniego et al., 1999; Alva-Valdivia et al., 2000). In particular, phylogenetic analyses have shown that the *Phaseolus* spp. can be grouped into two major sister clades: clade A, which comprises the Pauciflorus, Pedicellatus, and Tuerckheimii groups, and the weakly resolved species (i.e., *P. glabellus*, *P. macrolepis*, *P. microcarpus*, and *P. oaxacanus*); and clade B, which comprises the Filiformis, Vulgaris, Lunatus, Leptostachyus, and Polystachios groups (Delgado-Salinas et al., 2006). Thus, eight principal crown clades that show some morphological, ecological and bio-geographic distinctions characterize the *Phaseolus* genus, and their formation occurred relatively late on, with an average age of ~2 Ma ago



**FIGURE 1 | Geographic distribution of *Phaseolus* species richness in America.** The colors in the legend indicate the number of *Phaseolus* species present in a defined geographic area (modified from Ramírez-Villegas et al., 2010; permissions for reproduction have been obtained from the copyright holders).

(Delgado-Salinas et al., 2006). The oldest group is *Vulgaris*, which has been dated at  $\sim 4$  Ma ago (Delgado-Salinas et al., 2006).

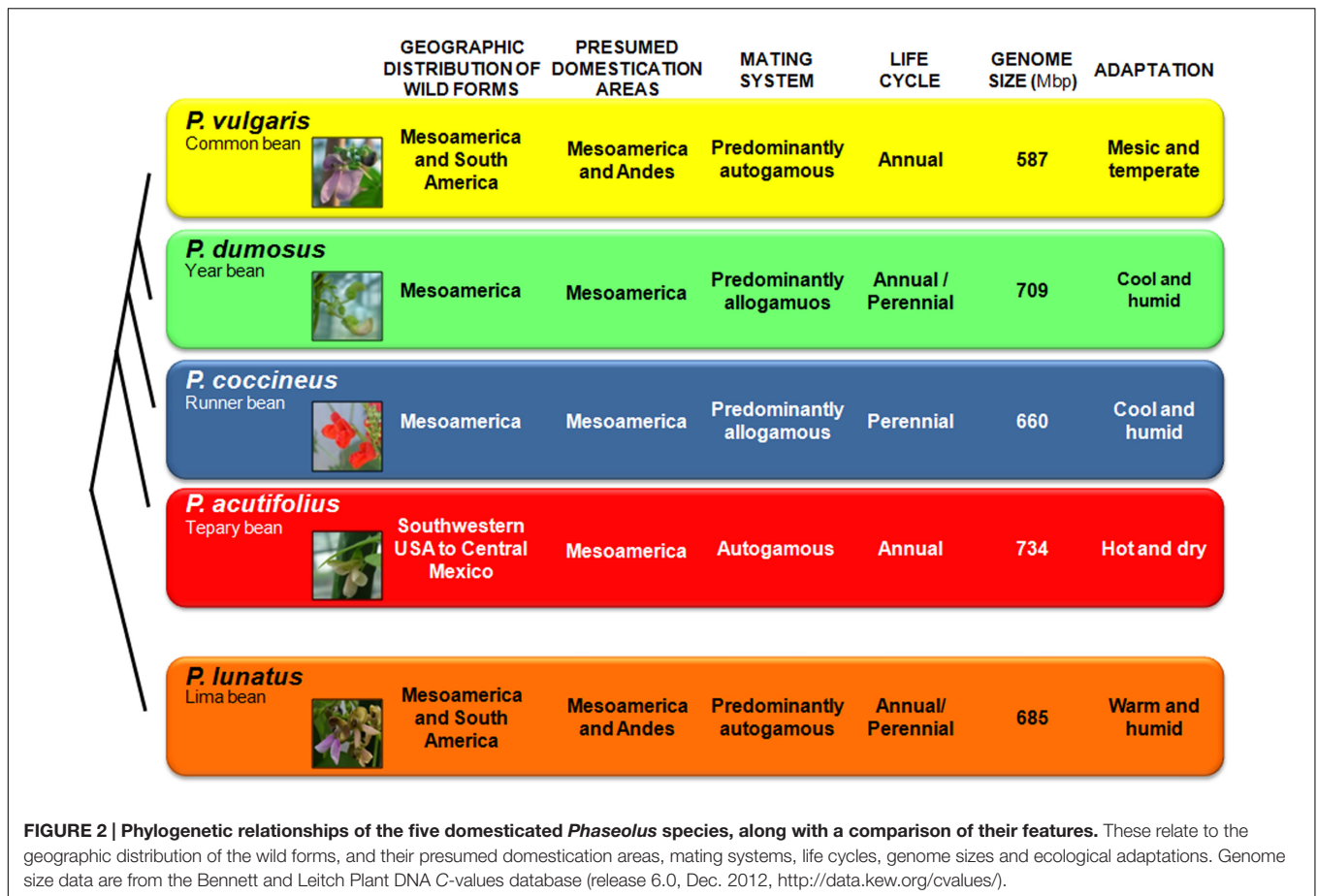
Five *Phaseolus* species were domesticated: the common bean *P. vulgaris*; the year bean *P. dumosus* Macfad.; the runner bean *P. coccineus* L.; the tepary bean *P. acutifolius* A. Gray; and the Lima bean *P. lunatus* L.. The Lima bean belongs to the *Lunatus* group, the formation of which has been dated at  $\sim 2$  Ma ago (Delgado-Salinas et al., 2006); it is a predominantly autogamous species that includes both annual determinate bush types and indeterminate climbers that are often perennials, due to their enlarged tap root (Baudoin, 1988; Salunkhe and Kadam, 1998) (Figure 2). All the other *Phaseolus* crop species (i.e., *P. vulgaris*, *P. dumosus*, *P. coccineus*, and *P. acutifolius*) belong to the *Vulgaris* group. In particular, *P. vulgaris*, *P. dumosus*, and *P. coccineus* are very closely related, and these three species are partially intercrossable, although only when *P. vulgaris* is the female parent (Mendel, 1866; Wall, 1970; Shii et al., 1982; Hucl and Scoles, 1985). This is despite their marked differences in mating systems and life cycles, as *P. vulgaris* is predominantly autogamous

and annual, *P. coccineus* is predominantly allogamous and perennial, and *P. dumosus* has intermediate characteristics between *P. coccineus* and *P. vulgaris* (Figure 2). Analysis of sequence data of the  $\alpha$ -amylase inhibitor gene indicated that *P. vulgaris* diverged from *P. dumosus* and *P. coccineus*  $\sim 2$  Ma ago (Gepts et al., 1999). *P. acutifolius*, is an annual species with a highly selfing reproductive system.

## GEOGRAPHIC DISTRIBUTION, ORIGIN AND ADAPTATION OF THE WILD FORMS OF DOMESTICATED *PHASEOLUS* SPECIES

### Geographic Distribution and Origin

The wild forms of *P. vulgaris* and *P. lunatus* are distributed in both Mesoamerica and South America, while those of *P. dumosus*, *P. coccineus*, and *P. acutifolius* have a geographic



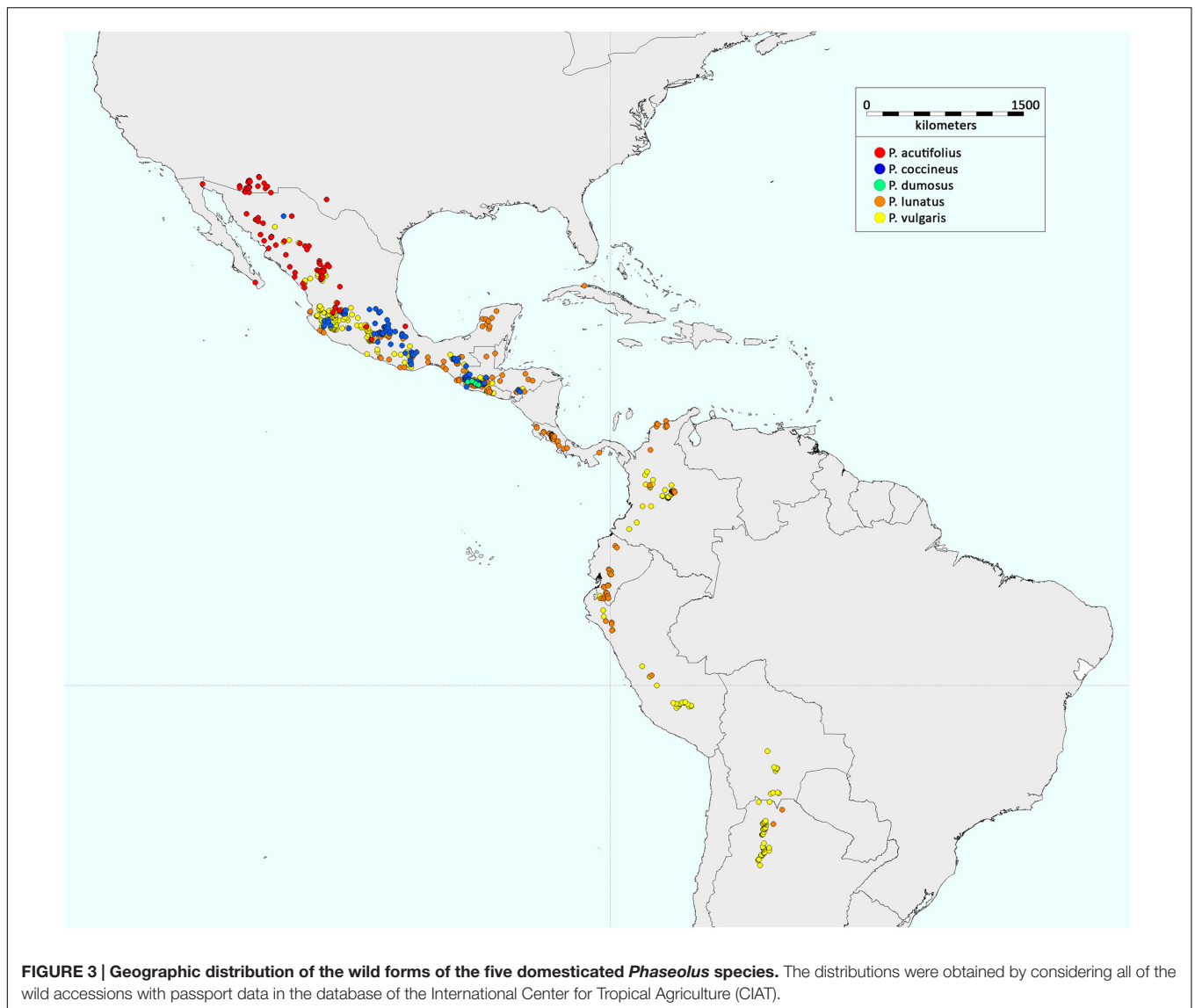
distribution that is restricted to Mesoamerica (Figure 3). There have been numerous studies that have investigated the origin and evolution of *P. vulgaris*, which among these five domesticated *Phaseolus* spp. has the greatest economic importance. In contrast, there have been very few studies into the origins of *P. lunatus*, *P. acutifolius*, *P. coccineus*, and *P. dumosus*.

### *Phaseolus vulgaris*

As proposed by Gepts (1988), a gene pool structure is identified by the observation within a biological species of strong population differentiation due to reproductive isolation, often associated to different adaptation, which can be due to geography and/or sexual incompatibility. Thus, the wild forms of the common bean, which grow from northern Mexico to northwestern Argentina (Toro et al., 1990; Figure 3), are characterized by three eco-geographic gene pools. Two of these, the Andean and Mesoamerican, are the major gene pools of the species, and they include both wild and domesticated forms (Bitocchi et al., 2013). The third gene pool is represented by wild populations that grow in northern Peru and Ecuador, in a narrow altitudinal fringe on the western and eastern slopes of the Cordillera, a region characterized by diverse environmental conditions that differ from those on which the other wild Andean forms, including Colombian populations, grow (Debouck et al., 1993). Several studies have indicated the specific patterns of allelic

frequencies and linkage disequilibrium that are characteristic of the populations from northern Peru and Ecuador (Papa and Gepts, 2003; Kwak and Gepts, 2009; Nanni et al., 2011; Bitocchi et al., 2012; Desiderio et al., 2013; Rodriguez et al., 2016; Rendón-Anaya et al., 2017). This gene pool has only been described for wild populations, with no domesticated forms ever found. Moreover, the populations from northern Peru and Ecuador are characterized by a specific phaseolin type, known as 'Inca' (I), which is not found in individuals outside of this geographic location (Kami et al., 1995). In particular, the findings of Kami et al. (1995) indicated that the Inca sequence of the portion of the gene that codes for the protein phaseolin is ancestral to the other types of phaseolin in individuals from the Mesoamerican and Andean major gene pools. They thus indicated northern Peru and Ecuador as the area of origin of the common bean, from where, subsequently, the species became widespread northwards (Mesoamerica) and southwards (Andes), hence leading to the formation of the two major gene pools that are characteristic of *P. vulgaris*.

However, more recently, this evolutionary scenario was called into question by the studies of Rossi et al. (2009), and in particular, of Bitocchi et al. (2012), where their data analysis clearly indicated that both the Andean and the Inca gene pools are derived from independent introductions from Mesoamerica (Figure 4). In contrast to the previous studies that were based on



multilocus molecular markers, Bitocchi et al. (2012) investigated the evolutionary history of the common bean using nucleotide sequences for five gene fragments. The first outcome was the clear confirmation of a bottleneck that occurred prior to domestication for the Andean gene pool. This was suggested previously by other studies (i.e., Rossi et al., 2009), although it was the lower mutation rate characteristic of nucleotide data compared to multilocus molecular markers that allowed the strong effect of the bottleneck on the genetic diversity of the Andean wild germplasm to be highlighted (i.e., a 90% reduction in diversity compared to the Mesoamerican wild gene pool; Bitocchi et al., 2012). Moreover, as sequence data from a single locus of a few 100 base pairs are less prone to recombination than multilocus molecular markers, these data allowed the identification of a strong population structure in Mesoamerica, although for the first time without any clear distinction between the Mesoamerican and Andean wild gene pools (Bitocchi et al., 2012). Indeed, the

phylogenetic relationships between the different groups showed two Mesoamerican groups (from Mexico) that appeared to be more closely related, one to the northern Peru–Ecuador gene pool and the other to the Andean gene pool. This led to the conclusion that each of these gene pools from South America originated through different migrations from the Mesoamerican populations of central Mexico. This hypothesis was also supported by subsequent studies (Schmutz et al., 2014; Rendón-Anaya et al., 2017). However, Rendón-Anaya et al. (2017), differently from Bitocchi et al. (2012) and Desiderio et al. (2013), by analyzing whole genome sequencing data from 18 *P. vulgaris* accessions (eight wild and two domesticated Mesoamerican accessions; one wild and two domesticated Andean accessions and five accessions from northern Peru and Ecuador), suggested that the northern Peru and Ecuador group should be considered as a sister species on the basis of the observation of a complete separation between the northern Peru and Ecuador genotypes



and a group composed by both Mesoamerican and Andean genotypes. However, from the works of Bitocchi et al. (2012) and Desiderio et al. (2013), presenting a larger sample compared to Rendón-Anaya et al. (2017), for both nuclear and chloroplast genome, wild populations from Mesoamerica closely related to the northern Peru and Ecuador group were identified, suggesting caution on the claim that the northern Peru and Ecuador group represents a different species. The whole genome sequencing analysis conducted by Schmutz et al. (2014), who estimated the divergence time between the Andean and Mesoamerican gene pools by applying demographic modeling, suggests that the wild populations in the Andes were derived from an ancestral Mesoamerican population ~165,000 years ago. The first attempt to date the split between the Andean and Mesoamerican gene pools was represented by the study of Gepts et al. (1999), where this event was estimated to have occurred ~500,000 years ago on the basis of the analysis of the  $\alpha$ -amylase inhibitor and internal transcribed spacer sequence data. Mamidi et al. (2013) identified an earlier date (~110,000 years ago) compared to that estimated by Schmutz et al. (2014). This recent divergence is in agreement with the high similarity between the genomes and with the recent observation by Vlasova et al. (2016) that most of the bean-specific gene family expansion predates the split between Mesoamerica and the Andes.

### The Other *Phaseolus* Crop Species

The *P. lunatus* wild forms are widely distributed from central Mexico to northern Argentina (Allard, 1960; Heiser, 1965;

Freytag and Debouck, 2002; **Figure 3**). Studies of the evolutionary history of *P. lunatus* have essentially been based on limited genomic data. In particular, most of these have relied on the analysis of two intergenic spacers of the chloroplast DNA (i.e., *atpB-rbcL*, *trnL-trnF*) and the sequence of the nuclear ribosomal 5.8S and flanking internal transcribed spacers (the ITS region) (Motta-Aldana et al., 2010; Serrano-Serrano et al., 2010, 2012; Andueza-Noh et al., 2013). Serrano-Serrano et al. (2010) analyzed these nuclear and non-coding chloroplast DNA markers in a collection of 59 wild Lima bean accessions and six related *Phaseolus* spp., three of which were of Andean distribution (i.e., *P. augusti*, *P. pachyrrhizoides*, and *P. bolivianus*), with the others distributed in Mesoamerica (i.e., *P. leptostachyus*, *P. marechalii*, and *P. novoleonensis*). Using neighbor-joining tree analysis, they identified three divergent wild Lima bean gene pools: Ecuador and northern Peru (AI); Mexico, and mainly the area to the west and northwest of the Isthmus of Tehuantepec (MI); and Mexico, and mainly the area to the east and southeast of the Isthmus of Tehuantepec (MII). Moreover, they suggested an Andean origin for wild Lima bean, as had been reported previously (Caicedo et al., 1999; Delgado-Salinas et al., 1999; Fofana et al., 1999). Their main evidence was the close phylogenetic relationship between the wild Lima bean and the related Andean *Phaseolus* species. In agreement with the study of Delgado-Salinas et al. (2006), Serrano-Serrano et al. (2010) indicated a relatively recent origin of *P. lunatus*, to during the Pleistocene and after the major Andean orogeny, ~2 to 5 Ma ago (Gregory-Wodzicki, 2000; Young et al., 2002). More recent studies based on the same markers and large samples of wild materials (Serrano-Serrano et al., 2012; Andueza-Noh et al., 2013) have confirmed the structure of the Mesoamerican wild populations, with the identification of the two main groups (MI, MII) with a geographic distribution that is probably related to adaptation to the different environments. Indeed, as indicated by Serrano-Serrano et al. (2012), the MI gene pool is mainly distributed in tropical dry forests over the Pacific coastal plain in Mexico at an average altitude of ~450 m a.s.l., with a small group of accessions on the western side of the Neo-Volcanic Axis at higher altitudes (from 1,250 to 1,810 m a.s.l.), while the MII gene pool is present in the Mexican lowlands (~550 m a.s.l.) along the Atlantic coast (Mexican gulf) and the Yucatan peninsula, and to the southeast of the Isthmus of Tehuantepec, the Caribbean, and South America.

Recently, Martínez-Castillo et al. (2014) analyzed 67 wild *P. lunatus* accessions from Mexico using 10 microsatellite (simple sequence repeats; SSR) markers. Through population structure analysis, they suggested that the genetic structure of the wild Lima bean in Mexico is more complex than previously thought, and they proposed three gene pools in Mesoamerica (MIa, MIb, and MII).

However, comparisons of the results of these studies show that there remain some disagreements, with different assignments to the diverse gene pools for the same accessions, according to the nuclear and chloroplast markers. Moreover, in some cases, the groups are not well supported statistically, and thus further studies are needed to more deeply understand the evolution of the Lima bean *P. lunatus*.

As mentioned above, the other three *Phaseolus* crop species (i.e., *P. acutifolius*, *P. coccineus*, and *P. dumosus*) are distributed in North and Central America. The wild forms of the tepary bean, which include *P. acutifolius* var. *acutifolius* and *P. acutifolius* var. *tenuifolius*, grow in the region from Central Mexico to southwestern USA (Blair et al., 2002; **Figure 3**). Indeed, this species is believed to have originated within this geographic range (Freeman, 1912, 1913; Nabhan and Felger, 1978; Manshardt and Waines, 1983).

The wild forms of the runner bean *P. coccineus* are distributed from Chihuahua (Mexico) to Panama (Delgado Salinas, 1988). The few archeological remains available (Delgado Salinas, 1988; Kaplan and Lynch, 1999) also indicate this area as the center of origin of *P. coccineus*.

For the wild form of the year bean *P. dumosus*, the geographic distribution is centered on a very narrow area in Guatemala, and this is where it has been suggested that *P. dumosus* originated (Schmit and Debouck, 1991).

## Adaptation

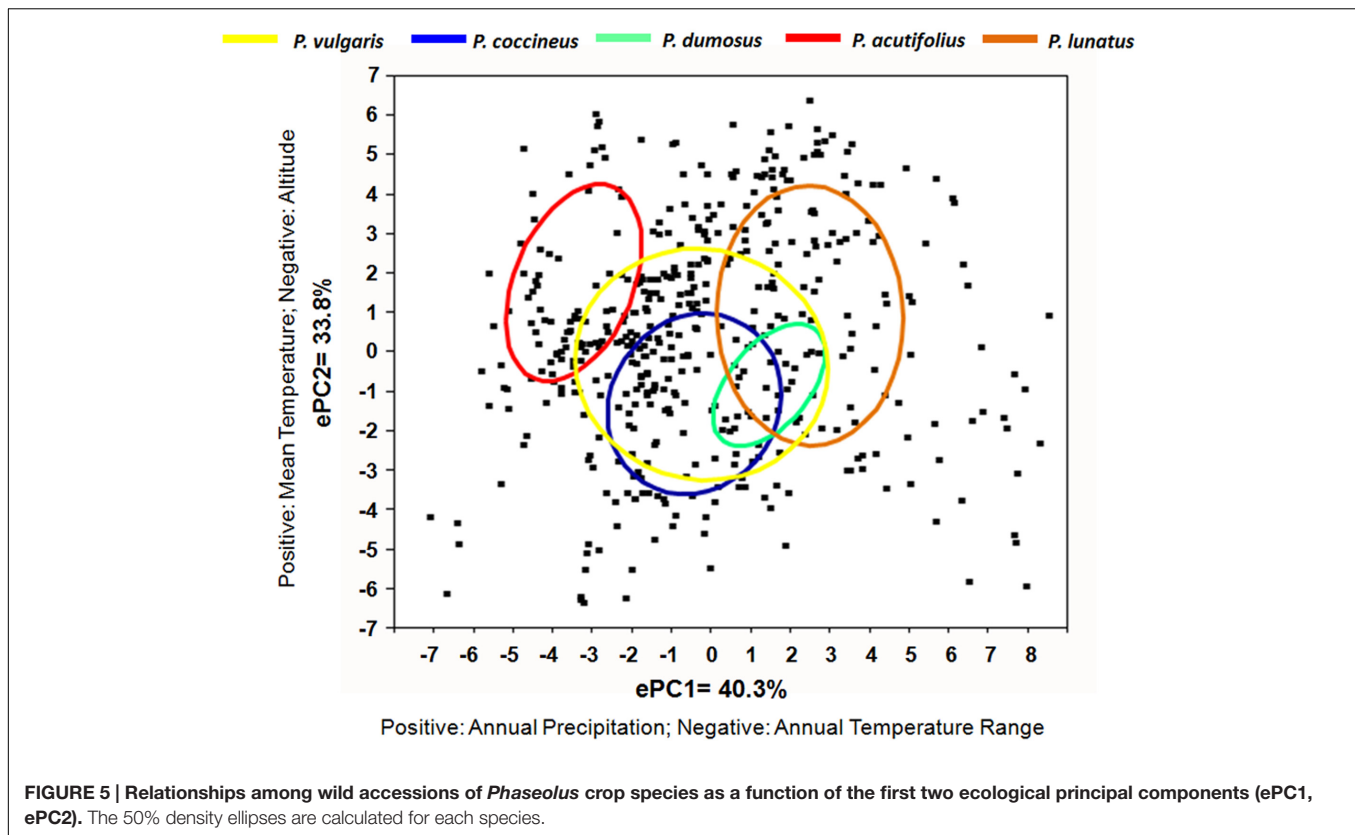
Genetic diversity is not uniformly distributed throughout the variety of climatic and environmental conditions under which a species grows. Consequently, it depends not only on demographic processes (e.g., genetic drift), but also on adaptation, in terms of the ability to cope with specific environmental conditions. These conditions can include extremes of cold and heat, lack or excess of water, different light intensities and duration, and diverse soil conditions and pest and disease pressures. As indicated above, the wild forms of *Phaseolus* crop species grow under a variety of different environmental and climatic conditions, and their geographic distribution mirrors their diverse patterns of adaptation to the different ecological niches, as well as their life histories and reproductive systems (**Figure 2**). *P. vulgaris* is adapted to warmer temperatures (mesic and temperate soil temperatures) at lower altitudes, with a rainfall of ~1,100 mm/year (Supplementary Table S1). *P. coccineus* is found in more humid environments, at cooler temperatures, and at higher altitudes, while *P. dumosus* is characterized by intermediate adaptation. *P. acutifolius* is a drought-tolerant species that originated in warmer and more arid environments (e.g., it is grown in the arid lands of Mexico and in southeast USA). Finally, *P. lunatus* is particularly suited to low-altitude humid and sub-humid climates, as well as warm temperate zones (**Figure 2** and Supplementary Table S1). These observations are confirmed by principal component analysis (**Figure 5**). This was calculated from the data on 20 ecological variables (Supplementary Table S1) using DIVA-GIS 7.5<sup>1</sup>, as inferred from the geographical coordinates of the collection sites of the wild accessions for the passport data present in the database of the International Centre for Tropical Agriculture (CIAT). Moreover, as can be seen in **Figure 5**, among these five species, *P. vulgaris* shows the widest ecological adaptation.

There is little in the literature that is aimed at highlighting the genetic basis of adaptation, especially for the wild forms of these crop species, although a few recent studies have been

focused on *P. vulgaris*. Rodriguez et al. (2016) carried out a study on the environmental adaptation of wild *P. vulgaris*. They investigated the role of demographic processes (e.g., genetic drift) and selection for adaptation in the shaping of the current genetic structure of wild *P. vulgaris*. This analysis was based on 131 single nucleotide polymorphisms (SNPs) on a sample of 417 wild *P. vulgaris* accessions that are representative of the geographic distribution of the wild forms of the species. They first investigated the spatial distribution of the genetic diversity of the wild forms of *P. vulgaris*, using a landscape genetics approach that was based on an individual centered analysis to avoid sampling bias (**Figure 6**). Briefly, they delineated a circular neighborhood of 100-km radius around each georeferenced accession, and the accessions falling within each neighborhood were used to calculate the relative unbiased gene diversity,  $H_e$  (Nei, 1978). The correlation between  $H_e$  and the neighborhood size was not significant ( $r = 0.040$ ,  $n = 299$ ,  $P = 0.482$ ), and the mean size of each neighborhood was 40.6 individuals and 83.3% of the neighborhoods included >10 individuals (Rodriguez et al., 2016). By using this approach, they observed high genetic diversity across Mexico, from the state of Oaxaca to Durango, with a depression in the genetic diversity in an area that lies approximately across the regions of Guerrero, Morelos, Puebla and Estado de Mexico. Possible explanations for the reduced diversity of this area in Mexico include natural selection due to an environment that is too arid for *P. vulgaris*, or a genetic bottleneck caused by the volcanic activities that were frequent in this area in ancient times (Márquez et al., 1999; Siebe, 2000; Siebe et al., 2004; Plunket and Uruñuela, 2012). Low diversity was also found in Guatemala, Costa Rica and Colombia, and particularly in the Honduras (**Figure 6**).

The approach applied by Rodriguez et al. (2016) also highlighted local adaptation at the continental level. The rationale behind the study was that the use of spatial data in combination with genetic diversity data would allow discrimination between the effects of geography and ecology; i.e., demographic processes vs. selection (Bradburd et al., 2013; Wang et al., 2013; Guillot et al., 2014; Kraft et al., 2014). As evidence of the effectiveness of this approach, it was possible to disentangle the effects of geography from ecology in the shaping of the genetic patterns observed, and correlations between markers and ecological variables were detected. By scanning the SNP markers used for the analyses, 26 loci (19.8%) were identified as under signatures of selection, seven of which (5.3%) showed strong probability levels. Although the proportion of loci under selection might be overestimated, as they were chosen among genes that were putatively involved in adaptation, different loci were found to have compatible functions with adaptation features, such as cold acclimation, chilling susceptibility, and mechanisms related to drought stress. Moreover, as well as the well-delineated genetic groups in the Mesoamerican gene pool, they demonstrated global structures for both the neutral loci and the loci under selection. Overall, these data suggest that the origin of the geographic structures might be the outcome of the expansion of the species and gene flow, including crop-to-wild introgression (Papa and Gepts, 2003; Papa et al., 2005), as was

<sup>1</sup><http://www.diva-gis.org/>



also recently confirmed by Rendón-Anaya et al. (2017) using whole genome sequencing analysis. Nonetheless, the adaptation to different environmental conditions might have led to the present population structure, with the subsequent limited long-range gene flow and divergent selection along an ecological cline of variation.

## DOMESTICATION OF *PHASEOLUS* SPP.

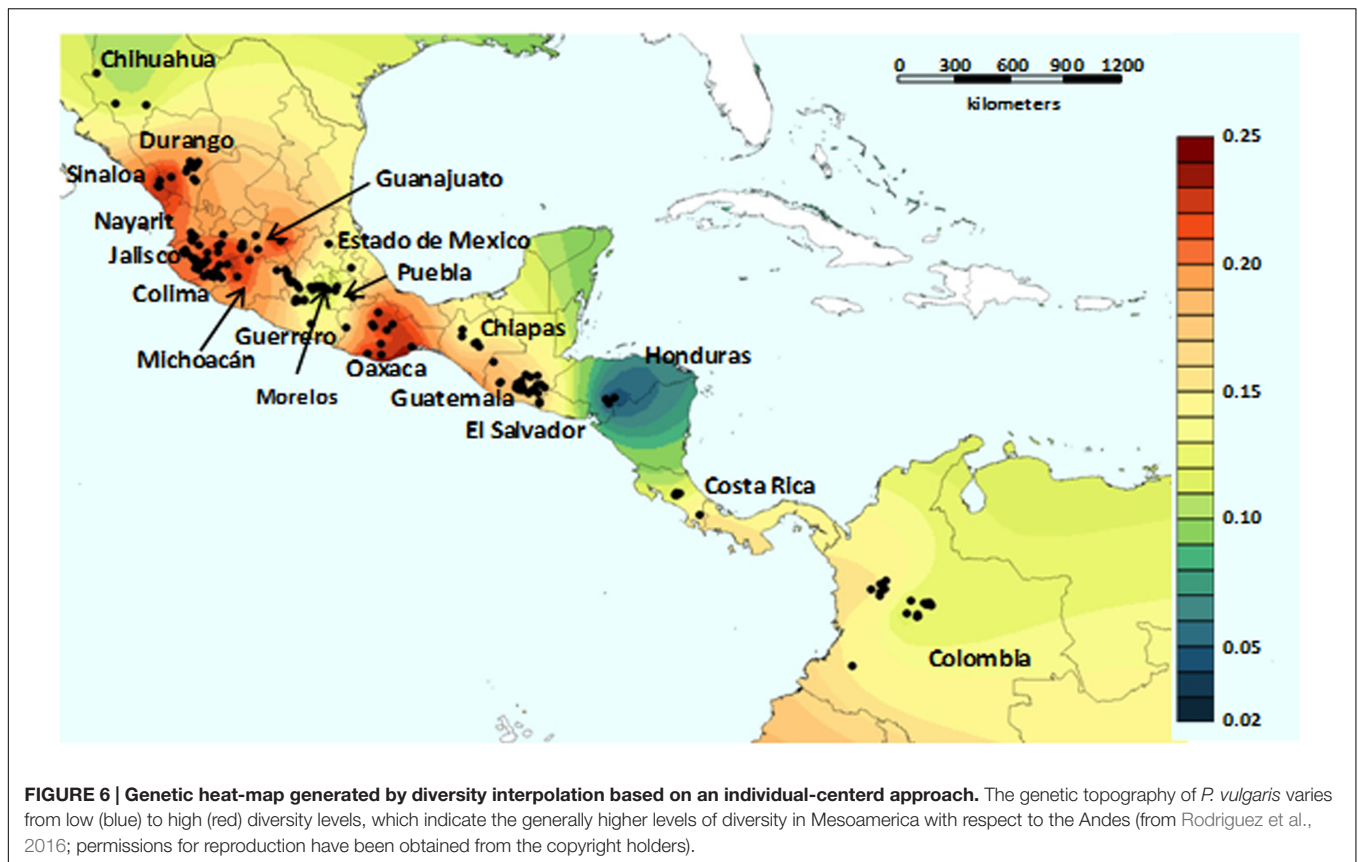
In the context of domestication, the *Phaseolus* spp. represent a unique model, where these five closely related species provide a unique example of multiple parallel domestication events, which allow to investigate this evolutionary process as a kind of replicate experiment. Two of these species (i.e., *P. vulgaris* and *P. lunatus*) have gone through at least two isolated and independent domestication processes, due to the distribution of the wild forms. For the common bean *P. vulgaris*, two independent events in the Americas (in Mesoamerica and the Andes) have been documented in several studies (for review, see Bellucci et al., 2014b), where the two major domesticated gene pools originated (Bitocchi et al., 2013). A similar scenario has been observed for *P. lunatus*, where at least two independent domestication events have been suggested. One of these was in the Andes, which appears to have given rise to the large-seeded landraces collectively known as the 'big Lima' cultivars (Motta-Aldana et al., 2010). The other was in Mesoamerica, which appears to have given rise to the great variety of small-seeded

Mesoamerican landraces (Motta-Aldana et al., 2010; Serrano-Serrano et al., 2012).

Not so much is known for the other three domesticated *Phaseolus* spp.. Through an analysis of a small set of wild and domesticated *P. coccineus* accessions and using chloroplast SSRs, Angioi et al. (2009a) identified two different wild genetic groups that paralleled the differentiation between two groups of the domesticated accessions, which suggested multiple domestication events for *P. coccineus* in Mesoamerica. A single domestication event was suggested for *P. dumosus* by Schmit and Debouck (1991) on the basis of seed protein data that were analyzed on a sample of 163 wild and domesticated accessions, as also for *P. acutifolius*, through studies based on phaseolin (Schinkel and Gepts, 1988), isoenzymes (Garvin and Weeden, 1994), SSRs (Blair et al., 2012), and SNPs (Gujaria-Verma et al., 2016). In particular, Gujaria-Verma et al. (2016) analyzed 645 SNPs markers on a wide sample that included both wild and domesticated *P. acutifolius* accessions, and they reported that domesticated tepary beans formed a tightly linked cluster that was subdivided into two major groups based on their eco-geographical origin (Central America and USA/Mexico). Moreover, the domesticated accessions were clearly separated from the wild, which suggested that it was likely that there had been an early domestication event that was followed by separation based on regions.

Overall, for the *Phaseolus* species, at least seven independent domestication events have occurred, with examples seen of both multiple and single domestication processes.





## Geographic Areas of *Phaseolus* ssp. Domestication

One of the major issues of evolutionary studies that are focused on the domestication of plant species is to identify the geographic areas where they were domesticated (Harlan, 1975).

### *Phaseolus vulgaris*

Considering the process of domestication within each gene pool, after the long debate for *P. vulgaris* on this issue (for review, see Bellucci et al., 2014b), recent studies have indicated that a single domestication event occurred within each gene pool (Mamidi et al., 2011; Nanni et al., 2011; Bitocchi et al., 2013).

For the Mesoamerican gene pool, several important crops were domesticated in Mexico, including maize, squash and common bean, as has been well documented in different studies based on both archeological and molecular data (for review, see Bitocchi et al., 2013). Recently, based on SSR data, Kwak et al. (2009) suggested the Rio Lerma–Rio Grande de Santiago basin in western-central Mexico as the putative geographic area where the common bean *P. vulgaris* was domesticated. The domestication area they suggested for the common bean did not overlap with that indicated for maize (central Balsas River drainage; Matsuoka et al., 2002; van Heerwaarden et al., 2011). Thus Kwak et al. (2009) proposed that maize and the common bean were probably domesticated in different regions and were instead reunited later in a single cropping system: the milpa, the cropping system that forms the basis of the Mesoamerican traditional

agriculture based on the intercropping of maize (*Zea mays* L.), the common bean, and squash (*Cucurbita* spp.). Bitocchi et al. (2013) used nucleotide data to suggest a different location for the domestication of the common bean: Oaxaca Valley, in the south of Mexico. This area does not coincide with that proposed by Kwak et al. (2009) for the common bean and that of maize (Matsuoka et al., 2002); however, it overlaps with one of the first areas in the spread of maize through human migration, along the Mexican rivers (Zizumbo-Villarreal and Colunga-GarcíaMarín, 2010).

The pinpointing of the geographic area where domestication took place for the common bean in the Andes has been more difficult, due to the low diversity characteristic of this germplasm. However, different studies have tried to identify the geographic area for this domestication. Chacón et al. (2005) analyzed the polymorphism at the level of the chloroplast DNA for a wide sample of common bean accessions from South America, and they suggested that central-southern Peru was the cradle of domestication of the Andean gene pool. Beebe et al. (2001) used amplified fragment length polymorphisms (AFLPs) to define a strict relationship between the domesticated forms and the wild beans from eastern Bolivia and northern Argentina, and suggested this location as the putative area of domestication, which was also supported recently by the study of Bitocchi et al. (2013).

More recently, on the basis of SNP data, the locations proposed by Bitocchi et al. (2013) for both the Mesoamerican and

Andean gene pools were confirmed by the study of Rodriguez et al. (2016). Here, Rodriguez et al. (2016) integrated the results obtained from spatial, phenotypic and molecular data with those from different disciplines, including archeological and glotto-chronological data, to pinpoint the domestication sites in Mesoamerica and the Andes. The low genetic distances between the wild forms and the domesticated forms indicated these as genetic groups located in the Oaxaca Valley in Mesoamerica and in a region from northern Argentina to southern Bolivia in the Andes, respectively. Consistent with these data, previous archeological data have indicated the early occurrence of domestication in these areas (Tarrago, 1980; Kaplan and Lynch, 1999). Recent glotto-chronological studies have also supported these conclusions, as within these areas the homeland sites of proto-languages for which ancient bean words were reconstructed showed times that are compatible with the domestication of *P. vulgaris* (Brown et al., 2014). We expect that future studies will better refine the areas of domestication. For this purpose will be very important also to conduct appropriate explorations. Indeed, as shown by Zizumbo-Villarreal et al. (2009) additional wild bean populations, not yet included in germplasm banks, can still be identified.

### The Other *Phaseolus* Crop Species

Similar to the common bean, the Lima bean *P. lunatus* has been domesticated at least twice in the Americas: once in the Andean region, and at another time in Mesoamerica. However, further studies are needed to investigate the domestication process of this species more deeply. Indeed, this still appears unclear especially for the Mesoamerican gene pool, where there remains open debate concerning its single or multiple domestication (Motta-Aldana et al., 2010; Serrano-Serrano et al., 2012; Andueza-Noh et al., 2013).

Motta-Aldana et al. (2010) analyzed chloroplast DNA and ITS polymorphisms in a sample of wild and landrace accessions of *P. lunatus*, through which they suggested one domestication event in the Andes of northwestern Peru and southern Ecuador, and a second in central-western Mexico, which they indicated as more likely to be in the area to the north and northwest of the Isthmus of Tehuantepec. Consistent with this study by Motta-Aldana et al. (2010), in their analysis of ITS data on a large sample of Mesoamerican wild and domesticated *P. lunatus*, Serrano-Serrano et al. (2012) proposed a single event of domestication in Mesoamerica. As indicated above, they showed evidence for two wild Mesoamerican gene pools with mostly contrasting geographic distributions. In their cluster analysis, all of the Mesoamerican landraces clustered together with the wild accessions from the MI gene pool, which is characteristic of central-western Mexico. This suggests a unique domestication event in an area of the states of Nayarit–Jalisco or Guerrero–Oaxaca, and not on the Peninsula of Yucatan, where *P. lunatus* is currently widespread and diverse.

Andueza-Noh et al. (2013) used two intergenic spacers of chloroplast DNA to confirm these Mesoamerican and Andean gene pools for *P. lunatus* (Gutiérrez-Salgado et al., 1995; Maquet et al., 1997; Lioi et al., 1998; Fofana et al., 2001; Motta-Aldana et al., 2010), and the two genetically and geographically distinct

groups within the Mesoamerican gene pool (MI, MII; Motta-Aldana et al., 2010; Serrano-Serrano et al., 2010, 2012). They pinpointed the domestication area for the Andean gene pool as mid-altitude western valleys between Peru and Ecuador in South America, as had already been suggested (Gutiérrez-Salgado et al., 1995; Fofana et al., 2001; Motta-Aldana et al., 2010). However, in contrast to the previous studies, Andueza-Noh et al. (2013, 2015) indicated multiple origins of domestication in Mesoamerica for *P. lunatus*. For the MI group, they indicated western-central Mexico as the domestication area, while they proposed a more restricted geographic area between Guatemala and Costa Rica for the MII group. However, they were aware that more studies involving more comprehensive geographic and genomic sampling are needed to define how the domestication processes and gene flow have shaped the current genetic structure of *P. lunatus* landraces.

For the other three *Phaseolus* crops (i.e., *P. acutifolius*, *P. coccineus*, and *P. dumosus*), nothing much is known about their domestication or where this process might have taken place. There have been very few studies on *P. acutifolius* and *P. coccineus*. For the tepary bean *P. acutifolius*, early studies based on phaseolin and isozyme analysis highlighted the controversy over the number of domestication events. Here, some studies proposed two domestication events in the northern and southern parts of the range (Manshardt and Waines, 1983), and others suggested a single origin but different locations for the domestication, as either in the Mexican state of Durango (Schinkel and Gepts, 1988) or the states of Sinaloa or Jalisco (Garvin and Weeden, 1994). The more recent study of Blair et al. (2012) was based on SSR data on a wide sample of wild and domesticated *P. acutifolius* accessions from its area of distribution. They indicated, as mentioned above, that a single domestication event was likely, and that the cultivars were most closely related to *P. acutifolius* var. *acutifolius* accessions from Sinaloa and northern Mexico.

*Phaseolus coccineus* is native to Mexico, Guatemala and Honduras (Delgado Salinas, 1988), and the wild forms are probably not all ancestral to the cultivated form. However, the area(s) where the domestication of *P. coccineus* took place are still not known. Spataro et al. (2011) used SSR data with a collection of wild and domesticated accessions, and they showed that most of the Mesoamerican landraces they examined closely resembled wild genotypes from Guatemala and Honduras, while only a few resembled wild Mexican forms. This would suggest that *P. coccineus* domestication either took place in that area, or that two domestication events took place (in Guatemala–Honduras and Mexico, separately) followed by extensive hybridisation with the cultivated forms from Guatemala and Honduras.

The distribution of wild *P. dumosus* (the year bean) is extremely narrow on the basis of findings to date, and it appears to be concentrated only in central southwestern Guatemala. Schmit and Debouck (1991) used phaseolin data together with information on vernacular names, and they reported that there is a single gene pool that was domesticated from a wild ancestor that is still present in Guatemala. Thus, they indicated a single domestication in Guatemala, and subsequent diffusion toward the humid highlands of Chiapas, Oaxaca, Puebla and

Veracruz in Mexico, and toward Costa Rica and the northern Andes.

## Domestication Bottleneck

The population genetics model of domestication predicts a reduction in diversity and increased divergence between wild and domesticated populations, due to demographic factors that affect the whole genome, and because of selection at target loci. Several interesting insights can be revealed by comparisons between different species (Glémin and Bataillon, 2009). There are many examples in the literature that have used different molecular markers and nucleotide data to show the reduction in genetic diversity of crop species compared with their wild progenitors (for review, see Bitocchi et al., 2013). Allopolyploid species, such as maize (*Z. mays*), are generally characterized by lower genetic bottleneck effects compared to autopolyploid species, such as the common bean *P. vulgaris*, even if other factors can have relevant roles, such as the life history (Bitocchi et al., 2013). Resequencing data have confirmed that in autopolyploid species, such as soybean (*Glycine max*) and rice (*Oryza sativa*, variety *japonica*) (Lam et al., 2010; Xu et al., 2012), reductions in diversity have arisen due to domestication, as also reported for silkworm and for mammalian species (Xia et al., 2009; vonHoldt et al., 2010; Lippold et al., 2011).

### *Phaseolus vulgaris*

For the common bean *P. vulgaris*, different studies have clearly identified a bottleneck due to domestication in both the Mesoamerican and Andean gene pools (e.g., Papa et al., 2005; Kwak and Gepts, 2009; Rossi et al., 2009; Mamidi et al., 2011; Nanni et al., 2011; Bitocchi et al., 2013, 2016a; Bellucci et al., 2014a). However, the reduction in diversity in the domesticated forms compared to the wild forms was greater in Mesoamerica compared to the Andes. Indeed, Bitocchi et al. (2013) reported that this loss of diversity was threefold greater for Mesoamerica compared to the Andes, and this was explained as the result of the bottleneck that occurred before domestication in the Andes. This thus strongly impoverished the genetic variability of the Andean wild germplasm, which led to a minor effect of the subsequent domestication bottleneck (i.e., sequential bottleneck). These outcomes demonstrate that the understanding of the level and the structure of genetic diversity of a species needs to be accompanied by a close appraisal of its evolutionary history.

Bellucci et al. (2014a) exploited next-generation sequencing technologies to analyze changes at the transcriptome level in *P. vulgaris* accessions from Mesoamerica, to investigate the domestication process in this gene pool more deeply. They used RNA sequencing technology and *de novo* transcriptome assembly to compare representative sets of wild and domesticated accessions of the common bean from Mesoamerica, and they reported the profound effects that domestication imposed on the genome variation and gene expression patterns of the common bean. Indeed, they showed that in addition to reduced nucleotide variation, the domesticated common bean showed reduced gene expression diversity, while in maize, the same reduction was not seen in parallel with reduced effects of domestication for nucleotide diversity (Hufford et al., 2012; Swanson-Wagner

et al., 2012). The expressed genomic regions lost half of the wild-bean nucleotide diversity during the domestication in Mesoamerica, and in parallel, the effects of domestication significantly decreased the diversity of gene expression (by 18%). For the first time, this demonstrated that loss of genetic variation has direct genome-wide phenotypic consequences on transcriptome diversity. The contigs identified as differentially expressed (in the comparison of domesticated vs. wild) were mostly down-regulated in the domesticated forms (by 74%). This indicated loss-of-function mutations (which are relatively frequent compared to gain-of-function changes) as a largely available source of variation that supports selection during rapid environmental changes (Olson, 1999). Such was the case for the transition from the wild to the cultivated agro-ecosystems. In support of this, as first noted by Darwin (1859), in domesticated plants, the domestication traits have a recessive genetic nature (Lester, 1989).

In addition to the case of differentially expressed genes, the genome-wide gene expression reported by Bellucci et al. (2014a) for the domesticated common bean *P. vulgaris* was on average lower than for the wild. They interpreted this result as the accumulation of slightly deleterious mutations due to hitchhiking (mostly loss-of-function, or with reduced expression) in *P. vulgaris*, and considered this as the ‘cost of domestication.’ This accumulation of loss-of-function (or reduced expression) mutations might also have been due to reduced effective recombination, which would have increased the frequency of deleterious mutations in the domesticated pool, and have had a negative influence on the fitness, as was suggested in rice (Lu et al., 2006).

### The Other *Phaseolus* Crop Species

In their analysis of chloroplast DNA and ITS polymorphisms in a sample of wild and landrace accessions of *P. lunatus* (the Lima bean), Motta-Aldana et al. (2010) observed a severe reduction in genetic diversity because of domestication in both the Mesoamerican and Andean gene pools (the MI wild accessions were used for co-mutations); in particular, the loss of diversity appeared stronger according to chloroplast DNA data (100%, 92.1%, for the Mesoamerican and Andean gene pools, respectively) than for ITS data (46.6%, 58.5%, respectively). This was confirmed for the Mesoamerican gene pool by Serrano-Serrano et al. (2012) and Andueza-Noh et al. (2013, 2015), through analysis of two intergenic spacers of chloroplast DNA (loss of diversity, 60.83%), SSR markers (loss of diversity, 44%), and the ITS region of ribosomal DNA (loss of diversity, 53%).

A bottleneck of domestication was also seen for *P. acutifolius*. Genetic diversity within the domesticated forms of the tepary bean is low, as has been shown by studies of phaseolin patterns (Schinkel and Gepts, 1988), isozymes (Schinkel and Gepts, 1989; Garvin and Weeden, 1994), AFLPs (Muñoz et al., 2006), and SSR markers (Blair et al., 2012). Considering the year bean *P. dumosus*, Schmit and Debouck (1991) analyzed phaseolin data and showed that the wild ancestral forms in central Guatemala show the highest diversity.

A different scenario has been reported in the few studies on domestication of the runner bean *P. coccineus*. Escalante et al.

(1994) indicated that the domestication process did not erode the genetic diversity of *P. coccineus*, and that the similar levels of genetic variation among the wild and cultivated materials were mainly due to the high gene flow between these two forms. This result was confirmed by Spataro et al. (2011) through an analysis of SSR data on a sample of wild and domesticated runner bean accessions.

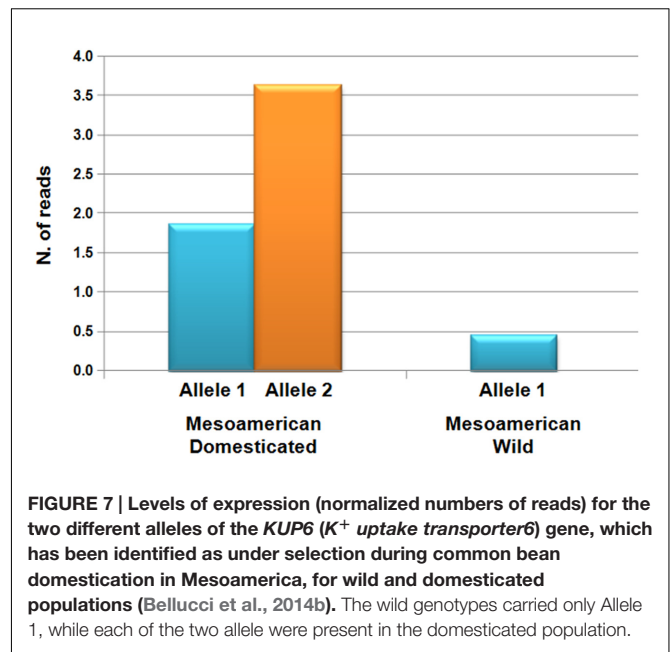
## Signatures of Selection during Domestication

Identification of the genes involved in the domestication process and knowledge of the regions of the genome where those genes are located and of the proportion of the genome affected by domestication are key to better exploitation of the diversity present in the wild relatives, and to enhance the achievements in breeding and crop improvement (Tanksley and McCouch, 1997; McCouch, 2004). As proposed by Cavalli-Sforza (1966) and Lewontin and Krakauer (1973), the identification of loci involved in adaptive processes can be obtained from population genetics expectations that predict that while drift has a homogeneous effect over the genome, selection is acting only for target loci and related linked loci due to the lack of recombination (hitchhiking). Thus, selected (and linked loci) are expected to depart from neutral expectation of diversity and divergence parameters. Moreover, as proposed by Papa et al. (2005), in gene flow between crops and wild forms, aberrant patterns of divergence and diversity can also be determined by the combined actions of asymmetric migration (Papa and Gepts, 2003) and selection at target loci.

Papa et al. (2007) used 2,506 AFLPs for a whole genome scan for the signature of selection due to domestication, and they estimated that about 16% of the genome of the common bean *P. vulgaris* appeared to be under the effects of selection. Bellucci et al. (2014a) used RNA sequencing technology, and after simulating the demographic dynamics during domestication, they reported that 9% of the genes were actively selected during domestication in Mesoamerica. Furthermore, in these contigs, selection induced a further reduction in the diversity of gene expression (by 26%), and was associated with a fivefold enrichment of the differentially expressed genes.

Bellucci et al. (2014a) also carried out a survey on the function of a subset of contigs that are putatively under selection, to determine whether they are known to be associated with the domestication process in other species, using either direct experimentation or through their function. Interestingly, among the genes putatively under selection that showed greater genetic diversity in the wild compared with the domesticated form, they found sequence homologs to: (i) genes that are involved in 'light' response pathway; e.g., *GIGANTEA* (*GI*), which has a pivotal role in the photoperiodic response (Mizoguchi et al., 2005; Hecht et al., 2007; Kim et al., 2012); (ii) genes that are pivotal to ensure correct hormonal perception, transport or biosynthesis; (iii) genes that are involved in seed development and traits; and (iv) genes that are involved in responses to environmental stress.

Another interesting example is the homolog of *YABBY5* (*YAB5*), which is a transcription factor that is implicated in



the regulation of seed shattering in cereal species, including sorghum (*Sorghum bicolor*), rice and maize (Lin et al., 2012). In most cases, Bellucci et al. (2014a) found evidence of positive selection associated with domestication, but in a few cases, this selection had increased the nucleotide diversity in the domesticated pool at a target locus associated with abiotic stress responses, flowering time, and morphology. In particular, for 2.8% of the genes putatively identified to be under the effect of selection by Bellucci et al. (2014a), there was no diversity in the wild forms, while there was diversity in the domesticated. They explained this as due to novel mutations (or standing variations) that were selected because of the crop expansion into new environments (diversifying selection) with unexpected biotic and abiotic stress, or because of selection for traits that improved the use of the plant organs by humans (de Alencar Figueiredo et al., 2008). An interesting example was given by the functional analysis of the drought- and growth-related (Osakabe et al., 2013) *KUP6* (*K<sup>+</sup> uptake transporter6*) gene, where this was significantly overexpressed in the domesticated compared to the wild (Figure 7), as if domestication had also increased the functional diversity of selected genes in addition to the increased nucleotide diversity.

Schmutz et al. (2014) also investigated whether their candidate genes were implicated in important domestication traits, such as flowering time and seed size. A total of 38 flowering genes were identified in the Mesoamerican and Andean candidate lists, while another subset of 15 genes was found to be associated with seed size in genome-wide association studies, and 11 genes contained SNPs that were associated with seed weight.

Recently, Bitocchi et al. (2016a) analyzed nucleotide sequences from a set of 49 gene fragments from a sample of 39 wild and domesticated Mesoamerican accessions of *P. vulgaris*. In this study, they applied the same approach as Bellucci et al. (2014a), and they identified several loci that showed signatures of

selection during common bean domestication in Mesoamerica. In particular, they had the possibility to see if their candidates were detected as outliers also in other studies of varying sizes, data types, and methodologies (Bellucci et al., 2014a; Schmutz et al., 2014; Rodriguez et al., 2016). They thus obtained independent evidence that four genes (i.e., *AN-Pv33*, *AN-Pv69*, *AN-DNAJ*, and *Leg223*) were targets of directional selection during common bean domestication. The gene function investigation for these genes highlighted that they are involved in plant resistance tolerance to biotic and abiotic stresses, such as heat, drought, and salinity. Moreover, another important outcome of Bitocchi et al. (2016a) was related to the observed excess of non-synonymous mutations in the domesticated germplasm. In particular, they observed a significantly higher frequency of polymorphisms in the coding regions compared to non-coding regions only in the domesticated beans. These mutations were mostly non-synonymous and were recently derived mutations present in genes related to responses to biotic and abiotic stresses. These data cannot be fully explained by the cost of domestication alone, but support a scenario where new functional mutations were selected for adaptation during domestication, showing that domestication also increased the functional diversity at target loci that enable the domesticated forms to successfully compete during the expansion and adaptation to new agro-ecological growing conditions.

## Phenotypic Convergent Evolution

The key aspect of domestication is the convergent phenotypic evolution that is associated with the adaptation to a novel agro-ecosystem, and to human needs. For instance, most domesticated animals were selected to maximize the production of useful products (e.g., meat, milk, and wool) and for their docile behavior, while crops were selected for the size of the plant organs used by humans (e.g., seeds and fruit) and for reduced, or lack of, seed dispersal. For these reasons, domestication provides us with a unique tool to understand the process of adaptation, to test evolutionary hypotheses, and to identify the molecular basis of phenotypic diversity. Several interesting insights can be revealed by comparisons among different species (Glémin and Bataillon, 2009), where, for instance, the population genetics model of domestication predicts a reduction in diversity and increased divergence between wild and domesticated populations due to demographic factors that affect the whole genome and because of selection at target loci.

In this regard, an example is given by the study of Nanni et al. (2011), in which wild and domesticated *P. vulgaris* accessions were analyzed by sequencing a genomic region of ~1,200 bp (*PvSHP1*) that is homologous to *SHATTERPROOF-1* (*SHP1*), a gene involved in the control of fruit shattering in *Arabidopsis thaliana* (Liljegren et al., 2000). The loss of fruit shattering has been under selection in most seed crops, to facilitate seed harvesting (Purugganan and Fuller, 2009), while in wild plants, this feature is a fundamental trait enabling seed dispersal. Expressed sequences that correspond to *SHP1* have also been identified in other species, such as, tomato, where it was indicated as having an important role in the regulation of both fleshy fruit expansion and the ripening process (Vrebalov et al., 2009),

which are together necessary to promote seed dispersal of fleshy fruit. In legumes, sequences orthologous to *Arabidopsis SHP1* have been identified in *Medicago*, pea and soybean (Hecht et al., 2005). Nanni et al. (2011) mapped *PvSHP1* on linkage group Pv06 of the common bean genome, and showed that it did not co-segregate with the *St* locus, which is responsible for the presence or absence of pod string, and was mapped on linkage group Pv02 using a domesticated (Midas) × wild (G12873) RIL population by Koinange et al. (1996). These results suggested that *PvSHP1* is not responsible for the observed phenotypic variation in *P. vulgaris* for fruit shattering. Similar results were found by Gioia et al. (2013b), who sequenced and mapped *PvIND* on the common bean genome, a sequence that is homologous to the *INDEHISCENT* gene (*IND*), which is the primary factory required for silique shattering in *A. thaliana*. *PvIND* mapped near the *St* locus; however, the lack of complete co-segregation between *PvIND* and *St* and the lack of polymorphisms at the *PvIND* locus correlated with the dehiscent/ indehiscent phenotype suggested that *PvIND* is not directly involved in pod shattering and is not the gene underlying the *St* locus (Gioia et al., 2013b).

Recently, Murgia et al. (2017) developed a phenotyping approach in *P. vulgaris* to evaluate the shattering syndrome in a segregating population. This is a promising approach for the identification of genetic factors that control the shattering trait in the common bean, and it will greatly facilitate comparative studies among legume crops, and also gene tagging.

Within the same species, the study of Schmutz et al. (2014) represents the first example of the possibility to investigate convergent evolution between the two gene pools of the common bean *P. vulgaris*. Indeed, a comparison of the results of selection in the two gene pools, in which independent domestications occurred, allowed them to determine whether to obtain the same convergent phenotypes, evolution took part in the selection of the same genomic regions or of completely different set of genes that code for the same phenotypes. Interestingly, only a small portion of the genome and of genes identified as putatively under selection during domestication were shared between the two gene pools, which suggested different genetic routes to domestication (Schmutz et al., 2014). This outcome appears to suggest that the sexually compatible Mesoamerican and Andean lineages with similar morphologies and life cycles underwent independent selection upon distinct sets of genes. However, taking into account that explicit demographic modeling was not used to generate an expectation of the number of potential false positive regions by Schmutz et al. (2014), another possible explanation for this result is that the lack of correlation between the two gene pools is due to a high level of false positives; i.e., regions of the genome with reduced diversity due to the stochastic effects of domestication bottlenecks. Regarding this consideration, Bitocchi et al. (2016a) compared their candidates for selection during domestication of common bean in Mesoamerica with those of other studies (Bellucci et al., 2014a; Schmutz et al., 2014; Rodriguez et al., 2016), and found that two (*AN-Pv69*, *AN-DNAJ*) out of the four strong candidates identified were detected as outliers by Schmutz et al. (2014)

only during Andean domestication. This implies that more studies are needed either to support or refute the lack of correlations between the two gene pools found by Schmutz et al. (2014).

## Dissemination of *Phaseolus* Crop Species Outside Their Centers of Origin

Beans are widely cultivated out of the Americas, and especially the common bean, which is the main *Phaseolus* crop species cultivated worldwide. For this reason, almost all the literature that has focused on investigation of the process of dissemination of beans out of their domestication centers is on *P. vulgaris*, with considerably less knowledge, if any, regarding the other *Phaseolus* crop species.

For common bean, a very complex scenario was highlighted by numerous studies, which includes: (i) several introductions from the New World, in combination with exchanges between continents, and among several countries within continents; (ii) new agro-ecological conditions experienced by this crop, implying new opportunities for both natural and human-mediated selection to act; and (iii) loss of spatial isolation characteristic of the Americas, which allowed hybridization and introgression between the Andean and Mesoamerican gene pools, and as a consequence, the occurrence of novel genotypes and phenotypes that transgressed the parental phenotypes for important agronomic and adaptive traits, such as, e.g., nutritional quality and resistance to biotic and abiotic stresses (Angioi et al., 2010; Blair et al., 2010; Gioia et al., 2013a).

## Patterns of Diversity of Beans Out of American Centers of Domestication

High levels of genetic diversity have been reported for common bean populations cultivated worldwide, and several continents and countries have been proposed as secondary centers of diversification for this species. These have included: the Iberian Peninsula (Santalla et al., 2002), the whole of Europe (Angioi et al., 2010, 2011; Gioia et al., 2013a), Brazil (Burle et al., 2010), central-eastern and southern Africa (Martin and Adams, 1987a,b; Asfaw et al., 2009; Blair et al., 2010), and China (Zhang et al., 2008).

In South America, a particular situation has emerged in Brazil. Although Brazil is closer to the Andes than to Mesoamerica, unexpectedly, it is the Mesoamerican *P. vulgaris* that is more prevalent (Burle et al., 2010). Multiple introductions of Mesoamerican germplasm in periods antecedent or successive to the discovery of the Americas might explain this pattern (Gepts et al., 1988).

In Africa, overall, the two gene pools are approximately equal in frequency, albeit there are strong differences between countries (Angioi et al., 2011; Okii et al., 2014). Such differences have been explained by the existence of at least partially independent seed networks in different countries (Asfaw et al., 2009), and because of selection due to dissimilar ecological and economic conditions among countries (Wortmann et al., 1998; Asfaw et al., 2009). Differential resistance to soil-borne diseases like

*Fusarium* root rot, and different yield performances arising from the 'interference' of improved genotypes released by national breeding programs have also been considered to explain the uneven distribution of the two gene pools across these regions (Blair et al., 2010).

In China, a prevalence of the Mesoamerican *P. vulgaris* has been observed (Zhang et al., 2008), although this was attributed mainly to founder effects (Zhang et al., 2008). The Himalayan region, as also the entire Indian subcontinent, shows high genetic diversity (Sofi et al., 2014; Rana et al., 2015).

Little is known about the dissemination of the common bean in India, albeit trading in the 16th century via the Red Sea and the Arabian Sea, and through the Hindustan Silk route, probably had a determinant role in the dissemination of this crop. It is also possible that the sea route discovered by the Portuguese explorer, Vasco da Gama, had a role in this dissemination. The genetic diversity in India also includes the combination of both the Mesoamerican and the Andean gene pools (Rana et al., 2015). Moreover, adaptation to micro-geographic conditions has been suggested for these landraces. Indeed, the analysis of >4000 landraces allowed the identification of several diverse clusters, irrespective of the place of collection, which also indicates a strong role for gene flow.

The Mesoamerican and Andean gene pools were both introduced into Europe. Studies carried out using several different marker types have shown that the Andean *P. vulgaris* predominates over Mesoamerican *P. vulgaris* (Gepts and Bliss, 1988; Lioi, 1989; Santalla et al., 2002; Logozzo et al., 2007; Angioi et al., 2010). The Andean type is largely predominant for the Iberian peninsula, Italy and central-northern Europe, where it is also prevalent on a local scale (Sicard et al., 2005; Angioi et al., 2009b). In the eastern part of Europe, the frequency of the Mesoamerican type tends to increase but it is always lower than the Andean (Papa et al., 2006). In their study of the expansion of the common bean *P. vulgaris* in Europe, Angioi et al. (2010) concluded that the intensity of the cytoplasmic bottleneck that resulted from this introduction into Europe was very low or absent (i.e., a loss of cpSSR diversity of ~2%).

Regarding the other *Phaseolus* crop species, there is little in the literature that has focused on the investigation of genetic diversity of *P. coccineus*, the allogamous sister species of *P. vulgaris*, in Europe. Some studies were conducted on small (Nowosielski et al., 2002; Sicard et al., 2005; Acampora et al., 2007; Boczkowska et al., 2012) and ample (Spataro et al., 2011; Rodriguez et al., 2013) spatial scales. The introduction of *P. coccineus* into Europe was probably contemporary with that of the common bean *P. vulgaris* (Westphal, 1974). Among the Mediterranean countries, *P. coccineus* is more widespread in Spain and Italy, while in northern Europe, it occurs more often in the UK and The Netherlands, where *P. coccineus* has often substituted for *P. vulgaris* (Santalla et al., 2004) as it is more adapted to cold temperatures and cool summers than *P. vulgaris* (Delgado Salinas, 1988; Rodino et al., 2007).

However, overall, *P. coccineus* is characterized by a narrower adaptability to environmental conditions than *P. vulgaris*. As previously mentioned, *P. coccineus* and *P. vulgaris* are cross-fertile when *P. vulgaris* is the maternal parent, and this might

have allowed hybridisation between these two species in Europe, where they often coexist in close sympatry in the same field. However, no evidence of introgression between common bean and runner bean has been found (Sicard et al., 2005). Data on nuclear and chloroplast variability of *P. coccineus* indicate that in Europe it has at least two main genetic groups (Spataro et al., 2011; Rodriguez et al., 2013). Of particular interest, there is a highly significant association between latitude and phenology for *P. coccineus* (Rodriguez et al., 2013). This relationship still holds when the effects of population structure for cpSSRs and nuSSRs is factored out. Therefore, this correlation is not just a consequence of the uneven geographic distribution of the two *P. coccineus* gene pools across Europe. It was then suggested that selection (probably for photoperiod sensitivity, and/or for low temperature), rather than migration and gene flow, has also had a role in shaping the population structure of *P. coccineus* in Europe (Rodriguez et al., 2013).

A comparison between Spanish and Mexican accessions of *P. vulgaris* and *P. coccineus* suggested that *P. coccineus* has maintained a high level of diversity since its introduction into Europe (Alvarez et al., 1998). A more recent study analyzed a worldwide collection of *P. coccineus*, and this analysis indicated that limited diversity of the runner bean *P. coccineus* appears to have been introduced into Europe, and that for nuclear markers, the European landraces show a reduction in diversity of 33% compared to that of the Mesoamerican landraces (Spataro et al., 2011). More recently, the use of chloroplast markers indicated a moderate-to-strong cytoplasmic bottleneck that followed the expansion of *P. coccineus* into Europe, with a reduction of 13% in chloroplast diversity (Rodriguez et al., 2013). As these markers are the same in number and type as those used by Angioi et al. (2010) to estimate the bottleneck of *P. vulgaris* following its expansion into Europe (2%), it can be concluded that the loss of diversity in *P. coccineus* appears to be stronger than in *P. vulgaris*.

Both nuclear and chloroplast analyses have shown that Mesoamerican and European *P. coccineus* accessions belong to distinct gene pools (Spataro et al., 2011; Rodriguez et al., 2013). It can be hypothesized that the differentiation of the European gene pool was due to adaptation to the new environment, and to genetic drift and a lack of introgression from wild forms.

The introduction of *P. lunatus* in Europe was very limited, with very few examples (Doria et al., 2012).

## A Role for Adaptive Introgression for the Evolution of European Beans?

Using isozyme markers, Santalla et al. (2002) estimated a high percentage of hybrids in their common bean landrace collection from the Iberian peninsula (25%). Then later, through the integration of cytoplasmic and nuclear analyses, Angioi et al. (2010) reported that about 44% of their wide collection of landraces that spanned almost all of the European countries appeared to be derived from at least one hybridisation event. In addition to the molecular results, Angioi et al. (2010) showed that seed size and coat traits vary with the level of introgression between the two gene pools. More recently, Gioia et al. (2013a) used population assignment techniques to also reveal extensive

hybridisation between the two gene pools in Europe, with a frequency of hybridisation that was almost fourfold greater in Europe (40.2%) than in the Americas (12.3%), which confirmed the findings of Angioi et al. (2010). This can be explained by the geographic isolation between the gene pools in the American centers of origin, and that following the introduction into Europe, genotypes from different gene pools often coexisted on very small spatial scales (i.e., in small cultivated areas), and thus had the chance to hybridize. Estimations of hybridisation in other parts of the world have always been <10%, and so much less than in Europe (Zhang et al., 2008; Asfaw et al., 2009; Blair et al., 2010; Burle et al., 2010). Taken all together, this evidence supports the hypothesis that the whole of Europe can be regarded as a secondary center of diversity for the common bean (Angioi et al., 2010), as also suggested by the work of Santalla et al. (2002) that was limited to the Iberian peninsula.

After its introduction into Europe, *P. vulgaris* was exposed to new and ample agro-ecological conditions. Thus, it is likely that any releases that were not adapted to the new conditions were initially purged by selection. Among the natural factors, biotic and abiotic stress were probably determinants in shaping the genetic structure of the European bean landraces. It is possible that their adaptation to long days, their cold tolerance, and their resistance to pests and diseases were crucial, and this would probably have led to a reduction in the diversity that was initially present in the founding populations. Additionally, the selection operated by farmers for seed color and size, and culinary, organoleptic and nutritional quality might also have had strong impact on the evolution of the European bean, as witnessed by the myriad of local bean populations with particular characteristics and specific names (Angioi et al., 2009b; Lioi and Piergiovanni, 2013; De Ron et al., 2016). Moreover, the documented scenario of extensive Mesoamerican × Andean gene pools through their hybridisation in Europe (Logozzo et al., 2007; Angioi et al., 2010; Gioia et al., 2013a) suggests that introgressive hybridisation might have been the fundamental ‘evolutionary stimulus’ (Anderson and Stebbins, 1954; Lewontin and Birch, 1966) that propelled and boosted bean evolution in Europe. Indeed, hybridisation can produce new genotypic and phenotypic combinations that do not occur in either of the parental taxa, and upon which selection might act.

Recently, Bitocchi et al. (2015) reported that European flint landraces grown *in situ* show adaptive introgression from modern maize. A key result of their study was that adaptation followed by hybridisation has been very rapid, with landraces capturing and increasing the frequency of favorable alleles over very short times (e.g., 50 years) (Bitocchi et al., 2015). This allows the hypothesis that this evolutionary mechanism might also have operated for the European bean landraces, which have a history that is some 10-fold longer. This is of interest not only for studies in evolutionary genetics, but also for plant breeders. Indeed, studies focused on hybridisation have shown the potential for the identification of functionally important regions of the genome (Arnold and Martin, 2010).

## The Possibility of Unraveling the Architecture of Adaptation of *Phaseolus vulgaris* in Europe

The potentiality of the studies at continental scale on the architecture of adaptation in common bean is confirmed by the possibility to apply a signature of selection approach using European landraces. Indeed, using methods for the identification of outlier loci for selection, Santalla et al. (2010) provided evidence that selective forces have had significant roles, particularly for seed size, growth habits, pest resistance and flowering time. This is intriguing, as selection for flowering time was probably a key element in the history of the bean also in its center of origin.

The BEAN\_ADAPT project<sup>2</sup> is funded through the 2nd ERA-CAPS call, ERA-NET for Coordinating Action in Plant Sciences. The main aim of this project is to dissect out the genetic basis and phenotypic consequences of the adaptation to new environments of *P. vulgaris* and its sister species *P. coccineus*, through the study of their introduction from their respective centers of domestication in the Americas and their expansion through Europe as a recent and historically well-defined event of rapid adaptation. BEAN\_ADAPT thus plans to characterize a large collection (11,500 accessions of each species) from three major genebanks by genotyping-by-sequencing. This will define the population structure and obtain subsets of genotypes for phenotyping (e.g., field, growth chamber) and for deeper genomic–transcriptomic–metabolomic characterisation. A multidisciplinary approach is planned (i.e., genomics, population/quantitative genetics, biochemistry, plant physiology) to unraveling the genetic bases of adaptation of these crops in new agro-ecological environments. The methods here rely on previous studies that have demonstrated the effectiveness of various analysis, such as the application of: (i) population genomics, to test for signatures of selection (e.g., Bellucci et al., 2014a); (ii) evolutionary metabolomics, which appear to be a very powerful approach to characterize molecular phenotypic changes due to domestication, and to identify traits under selection in wheat (Beleggia et al., 2016); (iii) association mapping studies of complex traits, such as flowering time (Raggi et al., 2014); (iv) analysis of signatures of selection by searching for ‘unusually high’ correlations between SNPs and environmental variables at the continental scale, as successfully applied by Rodriguez et al. (2016) for wild common bean from Mexico.

## CONCLUSION

This review presents a comprehensive overview of the current knowledge about the evolutionary history of the *Phaseolus* crop species. This takes us from the origins and evolution of their wild forms, with their co-evolution and interactions with humans and diverse environments during and after the domestication process, to their colonization of new environments

out of their centers of origin and domestication. The picture that has unraveled shows that the specific and almost unique features of the *Phaseolus* genus make it a very powerful model to address important evolutionary issues. In particular, among crops, *Phaseolus* represents a unique example of multiple parallel domestications for five closely related species, with two of these (i.e., *P. vulgaris* and *P. lunatus*) each domesticated independently in Mesoamerica and the Andes. This has resulted in at least a total of seven independent domestication events, involving species that have diverged relatively recently (~4 Ma; Delgado-Salinas et al., 2006) and that show similar genomic structures, making inter-specific comparisons feasible. This represents a ‘domestication experiment’ with full factorial design for three factors, as species, areas and wild/domesticated status, which can provide a deep understanding of the genomic architecture of domestication. Moreover, the study of different domestication events provides valuable replicates for the understanding of convergent evolution (i.e., different species or populations that evolved similar phenotypes) and its genomic determinants and effects. The possibility of considering these multiple parallel domestications as different replicates of the same experiment is also supported by the recent findings of Vlasova et al. (2016), who reported that both lineages of *P. vulgaris* and potentially all of the *Phaseolus* species share the same patterns of gene duplication that predate the divergence between the Mesoamerican and Andean gene pools.

Along with domestication, the complex patterns of dissemination of *Phaseolus* crops out of their centers of domestication represents a further key strength to unravel the genetic basis of plant adaptation. This is exactly what the BEAN\_ADAPT project is searching for: to dissect out the genetic basis and phenotypic consequences of the adaptation to new environments of the common bean and its sister species, the runner bean, through the study of their introduction and expansion through Europe, as a recent and historically well-defined event of rapid adaptation.

This approach can provide a model for future major environmental and socio-economic changes, such as increases in temperature, variability of rainfall, and new consumer preferences, which will be fruitful for both evolutionary biologists and plant breeders. For the evolutionary biologists, it will be of particular interest to compare the results obtained across different species and populations, to look for patterns of convergent evolution, at either the phenotypic or molecular level. Discovering genes and genetic mechanisms that contribute to phenotypic adaptation associated with environmental conditions and their mapping along the reference genome will provide a useful genetic tool for geneticists and breeders for the design of novel varieties.

## AUTHOR CONTRIBUTIONS

RP and EBi designed and wrote the manuscript. EBe, DS, LN, EBi, and RP contributed to the drafting of the manuscript especially for the sections on the origin and domestication of *Phaseolus* ssp.,

<sup>2</sup><http://www.beanadapt.org/>



while DR, MM, MR, GA and TG contributed to the drafting of the manuscript especially for the sections on dissemination out of the domestication centres of *Phaseolus* ssp.

## ACKNOWLEDGMENTS

This work was supported by grants from the ERA-NET for Coordinating Action in Plant Sciences-2nd ERA-CAPS call, BEAN\_ADAPT project, the Italian Government (MIUR; Grant number RBF13IDFM\_001, FIRB Project 2013) and

the Università Politecnica delle Marche (years 2015–2016). All of the appropriate permissions for reproduction of previously published Figures have been obtained from the respective copyright holders.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2017.00722/full#supplementary-material>

## REFERENCES

- Acampora, A., Ciaffi, M., De Pace, C., Paolacci, A. R., and Tanzarella, O. A. (2007). Pattern of variation for seed size traits and molecular markers in Italian germplasm of *Phaseolus coccineus* L. *Euphytica* 157, 69–82. doi: 10.1007/s10681-007-9397-3
- Allard, R. W. (1960). "Patterns of evolution in cultivated species," in *Principles of Plant Breeding*, ed. R. W. Allard (New York, NY: Wiley), 7–18.
- Alvarez, M. T., Sáenz de Miera, L. E., and Pérez de la Vega, M. (1998). Genetic variation in common and runner bean of the Northern Meseta in Spain. *Genet. Res. Crop Evol.* 45, 243–251. doi: 10.1023/A:1008604127933
- Alva-Valdivia, L. M., Goguitaichvili, A., Ferrari, L., Rosas-Elguera, J., Urrutia-Fucugauchi, J., and Zamorano-Orozco, J. J. (2000). Paleomagnetic data from the trans-mexican volcanic belt: implications for tectonics and volcanic stratigraphy. *Earth Planets Space* 52, 467–478. doi: 10.1186/BF03351651
- Anderson, E., and Stebbins, G. L. Jr. (1954). Hybridization as an evolutionary stimulus. *Evolution* 8, 378–388. doi: 10.2307/2405784
- Andueza-Noh, R. H., Martínez-Castillo, J., and Chacón-Sánchez, M. I. (2015). Domestication of small-seeded lima bean (*Phaseolus lunatus* L.) landraces in Mesoamerica: evidence from microsatellite markers. *Genetica* 143, 657–669. doi: 10.1007/s10709-015-9863-0
- Andueza-Noh, R. H., Serrano-Serrano, M. L., Chacón, M. I., Sánchez del-Pino, I., Camacho-Pérez, L., Coello-Coello, J., et al. (2013). Multiple domestications of the Mesoamerican gene pool of Lima bean (*Phaseolus lunatus* L.): evidence from chloroplast DNA sequences. *Genet. Res. Crop Evol.* 60, 1069–1086. doi: 10.1007/s10722-012-9904-9
- Angioi, S. A., Desiderio, F., Rau, D., Bitocchi, E., Attene, G., and Papa, R. (2009a). Development and use of chloroplast microsatellites in *Phaseolus* spp. and other legumes. *Plant Biol.* 11, 598–612. doi: 10.1111/j.1438-8677.2008.00143.x
- Angioi, S. A., Rau, D., Rodriguez, M., Logozzo, G., Desiderio, F., Papa, R., et al. (2009b). Nuclear and chloroplast microsatellite diversity in *Phaseolus vulgaris* L. from Sardinia (Italy). *Mol. Breed.* 23, 413–429. doi: 10.1007/s11032-008-9245-8
- Angioi, S. A., Rau, D., Attene, G., Nanni, L., Bellucci, E., Logozzo, G., et al. (2010). Beans in Europe: origin and structure of the European landraces of *Phaseolus vulgaris* L. *Theor. Appl. Genet.* 121, 829–843. doi: 10.1007/s00122-010-1353-2
- Angioi, S. A., Rau, D., Nanni, L., Bellucci, E., Papa, R., and Attene, G. (2011). The genetic make-up of the European landraces of the common bean. *Plant Genet. Res.* 9, 197–201. doi: 10.1017/S1479262111000190
- Arnold, M. L., and Martin, N. H. (2010). Hybrid fitness across time and habitats. *Trends Ecol. Evol.* 25, 530–536. doi: 10.1016/j.tree.2010.06.005
- Afaw, A., Blair, M. W., and Almekinders, C. (2009). Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) landraces from the East African highlands. *Theor. Appl. Genet.* 120, 1–12. doi: 10.1007/s00122-009-1154-7
- Baudoin, J. P. (1988). "Genetic resources, domestication and evolution of lima bean, *Phaseolus lunatus*," in *Genetic Resources of Phaseolus Beans*, ed. P. Gepts (Berlin: Springer), 393–407.
- Beebe, S., Rengifo, J., Gaitan, E., Duque, M. C., and Tohme, J. (2001). Diversity and origin of Andean landraces of common bean. *Crop Sci.* 41, 854–862. doi: 10.2135/cropsci2001.413854x
- Beleggia, R., Rau, D., Laidò, G., Platani, C., Nigro, F., Fragasso, M., et al. (2016). Evolutionary metabolomics reveals domestication-associated changes in tetraploid wheat kernels. *Mol. Biol. Evol.* 33, 1740–1753. doi: 10.1093/molbev/msw050
- Bellucci, E., Bitocchi, E., Ferrarini, A., Benazzo, A., Biagetti, E., Klie, S., et al. (2014a). Decreased nucleotide and expression diversity and modified coexpression patterns characterize domestication in the common bean. *Plant Cell* 26, 1901–1912.
- Bellucci, E., Bitocchi, E., Rau, D., Rodriguez, M., Biagetti, E., Giardini, A., et al. (2014b). "Genomics of origin, domestication and evolution of *Phaseolus vulgaris*," in *Genomics of Plant Genetic Resources*, eds R. Tuberosa, A. Graner, and E. Frison (Berlin: Springer), 483–507. doi: 10.1007/978-94-007-7572-5\_20
- Bitocchi, E., Bellucci, E., Giardini, A., Rau, D., Rodriguez, M., Biagetti, E., et al. (2013). Molecular analysis of the parallel domestication of the common bean (*Phaseolus vulgaris*) in Mesoamerica and the Andes. *New Phytol.* 197, 300–313. doi: 10.1111/j.1469-8137.2012.04377.x
- Bitocchi, E., Bellucci, E., Rau, D., Albertini, E., Veronesi, F., Attene, G., et al. (2015). European flint landraces grown in situ reveal adaptive introgression from modern maize. *PLoS ONE* 10:e0121381. doi: 10.1371/journal.pone.0121381
- Bitocchi, E., Nanni, L., Bellucci, E., Rossi, M., Giardini, A., Spagnoletti Zeuli, P. S., et al. (2012). Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data. *Proc. Natl. Acad. Sci. U.S.A.* 109, E788–E796. doi: 10.1073/pnas.1108973109
- Bitocchi, E., Rau, D., Benazzo, A., Bellucci, E., Goretti, D., Biagetti, E., et al. (2016a). High level of nonsynonymous changes in common bean suggests that selection under domestication increased functional diversity at target traits. *Front. Plant Sci.* 7:2005. doi: 10.3389/fpls.2016.02005
- Bitocchi, E., Rau, D., Rodriguez, M., and Murgia, M. L. (2016b). "Crop improvement of *Phaseolus* spp. through interspecific and intraspecific hybridization," in *Polyploidy and Hybridization for Crop Improvement*, ed. A. Mason (Boca Raton, FL: CRC Press), 218–280. doi: 10.1201/9781315369259-11
- Blair, M. W., González, L. F., Kimani, M., and Butare, L. (2010). Genetic diversity, inter-gene pool introgression and nutritional quality of common beans (*Phaseolus vulgaris* L.) from Central Africa. *Theor. Appl. Genet.* 121, 237–248. doi: 10.1007/s00122-010-1305-x
- Blair, M. W., Muñoz, L. C., and Debouck, D. G. (2002). Tepary beans (*P. acutifolius*): molecular analysis of a forgotten genetic resource for dry land agriculture. *Grain Legumes* 36, 25–26.
- Blair, M. W., Pantoja, W., and Muñoz, L. C. (2012). First use of microsatellite markers in a large collection of cultivated and wild accessions of tepary bean (*Phaseolus acutifolius* A. Gray). *Theor. Appl. Genet.* 125, 1137–1147. doi: 10.1007/s00122-012-1900-0
- Boczkowska, M., Bulinska-Radomska, Z., and Nowosielski, J. (2012). AFLP analysis of genetic diversity in five accessions of Polish runner bean (*Phaseolus coccineus* L.). *Genet. Res. Crop Evol.* 59, 473–478. doi: 10.1007/s10722-012-9798-6
- Bonifácio, E. M., Fonsêca, A., Almeida, C., Santos, K. G. B., and Pedrosa-Harand, A. (2012). Comparative cytogenetic mapping between the lima bean (*Phaseolus lunatus* L.) and the common bean (*P. vulgaris* L.). *Theor. Appl. Genet.* 124, 1513–1520. doi: 10.1007/s00122-012-1806-x
- Bradburd, G. S., Ralph, P. L., and Coop, G. M. (2013). Disentangling the effects of geographic and ecological isolation on genetic differentiation. *Evolution* 67, 3258–3273. doi: 10.1111/evo.12193
- Broughton, W. J., Hernández, G., Blair, M., Beebe, S., Gepts, P., and Vanderleyden, J. (2003). Beans (*Phaseolus* spp.): model food legumes. *Plant Soil* 252, 55–128. doi: 10.1023/A:1024146710611

- Brown, C. H., Clement, C. R., Epps, P., Luedeling, E., and Wichmann, S. (2014). The paleobiolinguistics of the common bean (*Phaseolus vulgaris* L.). *Ethnobiol. Lett.* 5, 104–115. doi: 10.14237/eb1.5.2014.203
- Burle, M. L., Fonseca, J. R., Kami, J. A., and Gepts, P. (2010). Microsatellite diversity and genetic structure among common bean (*Phaseolus vulgaris* L.) landraces in Brazil, a secondary center of diversity. *Theor. Appl. Genet.* 121, 801–813. doi: 10.1007/s00122-010-1350-5
- Caicedo, A. L., Gaitán, E., Duque, M. C., Toro Chica, O., Debouck, D. G., and Thome, J. (1999). AFLP fingerprinting of *Phaseolus lunatus* L. and related wild species from South America. *Crop Sci.* 39, 1497–1507. doi: 10.2135/cropsci1999.3951497x
- Cavalli-Sforza, L. (1966). Population structure and human evolution. *Proc. R. Soc. Lond. B Biol. Sci.* 164, 362–379. doi: 10.1098/rspb.1966.0038
- Chacón, S. M. I., Pickersgill, B., and Debouck, D. G. (2005). Domestication patterns in common bean (*Phaseolus vulgaris* L.) and the origin of the Mesoamerican and Andean cultivated races. *Theor. Appl. Genet.* 110, 432–444. doi: 10.1007/s00122-004-1842-2
- Choi, J. Y., Platts, A. E., Fuller, D. Q., Hsing, Y. I., Wing, R. A., and Purugganan, M. D. (2017). The rice paradox: multiple origins but single domestication in Asian rice. *Mol. Biol. Evol.* 34, 969–979. doi: 10.1093/molbev/msx049
- Darwin, C. (1859). *On the Origins of Species by Means of Natural Selection*. London: John Murray.
- de Alencar Figueiredo, L. F., Calatayud, C., Dupuits, C., Billot, C., Rami, J. F., Brunel, D., et al. (2008). Phylogeographic evidence of crop neodiversity in sorghum. *Genetics* 179, 997–1008. doi: 10.1534/genetics.108.087312
- De Ron, A. M., González, A. M., Rodiño, A. P., Santalla, M., Godoy, L., and Papa, R. (2016). History of the common bean crop: its evolution beyond its areas of origin and domestication. *ARBOR* 192, a317. doi: 10.3989/arb.2016.779n3007
- Debouck, D. G., Toro, O., Paredes, O. M., Johnson, W. C., and Gepts, P. (1993). Genetic diversity and ecological distribution of *Phaseolus vulgaris* in northwestern South America. *Econ. Bot.* 47, 408–423. doi: 10.1007/BF02907356
- Delgado Salinas, A. (1988). “Variation, taxonomy, domestication and germplasm potentialities in *Phaseolus coccineus*,” in *Genetic Resources of Phaseolus Beans*, ed. P. Gepts (Berlin: Springer), 441–463.
- Delgado-Salinas, A., Bibler, R., and Lavin, M. (2006). Phylogeny of the genus *Phaseolus* (Leguminosae): a recent diversification in an ancient landscape. *Syst. Bot.* 31, 779–791. doi: 10.1600/036364406779695960
- Delgado-Salinas, A., Turley, T., Richman, A., and Lavin, M. (1999). Phylogenetic analysis of the cultivated and wild species of *Phaseolus* (Fabaceae). *Syst. Bot.* 24, 438–460. doi: 10.2307/2419699
- Desiderio, F., Bitocchi, E., Bellucci, E., Rau, D., Rodriguez, M., Attene, G., et al. (2013). Chloroplast microsatellite diversity in *Phaseolus vulgaris*. *Front. Plant Sci.* 3:312. doi: 10.3389/fpls.2012.00312
- Doria, E., Campion, B., Sparvoli, F., Tava, A., and Nielsen, E. (2012). Anti-nutrient components and metabolites with health implications in seeds of 10 common bean (*Phaseolus vulgaris* L. and *Phaseolus lunatus* L.) landraces cultivated in southern Italy. *J. Food Compos. Anal.* 26, 72–80. doi: 10.1016/j.jfca.2012.03.005
- Escalante, A., Coello, G., Eguarte, L. E., and Piñero, D. (1994). Genetic structure and mating systems in wild and cultivated populations of *Phaseolus coccineus* and *P. vulgaris* (Fabaceae). *Am. J. Bot.* 81, 1096–1103. doi: 10.2307/2445471
- Fofana, B., Baudoin, J. P., Vekemans, X., Debouck, D. G., and du Jardin, P. (1999). Molecular evidence for an Andean origin and a secondary gene pool for the Lima bean (*Phaseolus lunatus*) using chloroplast DNA. *Theor. Appl. Genet.* 98, 202–212. doi: 10.1007/s001220051059
- Fofana, B., du Jardin, P., and Baudoin, J. P. (2001). Genetic diversity in the Lima bean (*Phaseolus lunatus* L.) as revealed by chloroplast DNA (cpDNA) variations. *Genet. Res. Crop Evol.* 48, 437–445. doi: 10.1023/A:1012016030318
- Fonseca, A., and Pedrosa-Harand, A. (2013). Karyotype stability in the genus *Phaseolus* evidenced by the comparative mapping of the wild species *Phaseolus microcarpus*. *Genome* 56, 335–343. doi: 10.1139/gen-2013-2025
- Freeman, G. F. (1912). *Southwestern Beans and Teparies*, Vol. 68. Tucson, AZ: University of Arizona.
- Freeman, G. F. (1913). The tepary, a new cultivated legume from the Southwest. *Bot. Gaz.* 56, 395–417. doi: 10.1086/331183
- Freytag, G. F., and Debouck, D. G. (2002). *Taxonomy, Distribution, and Ecology of the Genus Phaseolus (Leguminosae-Papilionoideae) in North America, Mexico and Central America*. Fort Worth, TX: Botanical Research Institute Of Texas.
- Garvin, D. F., and Weeden, N. F. (1994). Isozyme evidence supporting a single geographic origin for domesticated tepary bean. *Crop Sci.* 34, 1390–1395. doi: 10.2135/cropsci1994.0011183X003400050045x
- Gepts, P. (1988). “A Middle American and an Andean common bean gene pool,” in *Genetic Resources of Phaseolus Beans*, ed. P. Gepts (Dordrecht: Springer), 375–390. doi: 10.1007/978-94-009-2786-5\_15
- Gepts, P., and Bliss, F. A. (1988). Dissemination pathways of common bean (*Phaseolus vulgaris*, Fabaceae) deduced from phaseolin electrophoretic variability. II Europe and Africa. *Econ. Bot.* 42, 86–104. doi: 10.1007/BF02859038
- Gepts, P., Kmiecik, K., Pereira, P., and Bliss, F. A. (1988). Dissemination pathways of common bean (*Phaseolus vulgaris*, Fabaceae) deduced from phaseolin electrophoretic variability. I. The Americas. *Econ. Bot.* 42, 73–85. doi: 10.1007/BF02859036
- Gepts, P., Papa, R., Coulibaly, S., Mejía, A. G., and Pasquet, R. (1999). “Wild legume diversity and domestication - insights from molecular methods,” in *Wild Legumes Proceedings 7th MAFF International Workshop on Genetic Resources*, ed. D. Vaughan (Tsukuba: National Institute of Agrobiological Resources), 19–31.
- Gioia, T., Logozzo, G., Attene, G., Bellucci, E., Benedetelli, S., Negri, V., et al. (2013a). Evidence for introduction bottleneck and extensive inter-gene pool (Mesoamerica x Andes) hybridization in the European common bean (*Phaseolus vulgaris* L.) germplasm. *PLoS ONE* 8:e75974. doi: 10.1371/journal.pone.0075974
- Gioia, T., Logozzo, G., Kami, J., Spagnoletti Zeuli, P., and Gepts, P. (2013b). Identification and characterization of a homologue to the *Arabidopsis INDEHISCENT* gene in common bean. *J. Hered.* 104, 273–286. doi: 10.1093/jhered/ess102
- Glémin, S., and Bataillon, T. (2009). A comparative view of the evolution of grasses under domestication. *New Phytol.* 183, 273–290. doi: 10.1111/j.1469-8137.2009.02884.x
- Gregory-Wodzicki, K. M. (2000). Uplift history of the central and northern Andes: a review. *Geol. Soc. Am. Bull.* 112, 1091–1105. doi: 10.1130/0016-7606
- Guillot, G., Vitalis, R., le Rouzic, A., and Gautier, M. (2014). Detecting correlation between allele frequencies and environmental variables as a signature of selection. A fast computational approach for genome-wide studies. *Spat. Stat.* 8, 145–155. doi: 10.1016/j.spa.2013.08.001
- Gujaria-Verma, N., Ramsay, L., Sharpe, A. G., Sanderson, L. A., Debouck, D. G., Taran, B., et al. (2016). Gene-based SNP discovery in tepary bean (*Phaseolus acutifolius*) and common bean (*P. vulgaris*) for diversity analysis and comparative mapping. *BMC Genomics* 17:239. doi: 10.1186/s12864-016-2499-3
- Gutiérrez-Salgado, A., Gepts, P., and Debouck, D. G. (1995). Evidence for two gene pools of the Lima beans, *Phaseolus lunatus* L., in the Americas. *Genet. Res. Crop Evol.* 42, 15–28. doi: 10.1007/BF02310680
- Harlan, J. R. (1975). *Crops and Man*. Madison, WI: American Society of Agronomy.
- Hecht, V., Foucher, F., Ferrandiz, C., Macknight, R., Navarro, C., Morin, J., et al. (2005). Conservation of *Arabidopsis* flowering genes in model legumes. *Plant Physiol.* 137, 1420–1434. doi: 10.1104/pp.104.057018
- Hecht, V., Knowles, C. L., Vander Schoor, J. K., Liew, L. C., Jones, S. E., Lambert, M. J., et al. (2007). Pea LATE BLOOMER1 is a GIGANTEA ortholog with roles in photoperiodic flowering, deetiolation, and transcriptional regulation of circadian clock gene homologs. *Plant Physiol.* 144, 648–661. doi: 10.1104/pp.107.096818
- Heiser, C. B. (1965). Cultivated plants and cultural diffusion in nuclear America. *Am. Anthropol.* 67, 930–949. doi: 10.1525/aa.1965.67.4.02a00040
- Hucl, P., and Scoles, G. J. (1985). Interspecific hybridization in the common bean: a review. *HortScience* 20, 352–357.
- Hufford, M. B., Xu, X., van Heerwaarden, J., Pühjäärvi, T., Chia, J. M., Cartwright, R. A., et al. (2012). Comparative population genomics of maize domestication and improvement. *Nat. Genet.* 44, 808–811. doi: 10.1038/ng.2309
- Kami, J., Velásquez, V. B., Debouck, D. G., and Gepts, P. (1995). Identification of presumed ancestral DNA sequences of phaseolin in *Phaseolus vulgaris*. *Proc. Natl. Acad. Sci. U.S.A.* 92, 1101–1104. doi: 10.1073/pnas.92.4.1101
- Kaplan, L., and Lynch, T. (1999). *Phaseolus* (Fabaceae) in archaeology: AMS radiocarbon dates and their significance for pre-Columbian agriculture. *Econ. Bot.* 53, 261–272. doi: 10.1007/BF02866636
- Kim, Y., Yeom, M., Kim, H., Lim, J., Koo, H. J., Hwang, D., et al. (2012). GIGANTEA and EARLY FLOWERING 4 in *Arabidopsis* exhibit differential

- phase-specific genetic influences over a diurnal cycle. *Mol. Plant* 5, 152–161. doi: 10.1093/mp/sss005
- Koinange, E. M. K., Singh, S. P., and Gepts, P. (1996). Genetic control of the domestication syndrome in common bean. *Crop Sci.* 36, 1037–1045. doi: 10.2135/cropsci1996.0011183X003600040037x
- Kraft, K. H., Brown, C. H., Nabhan, G. P., Luedeling, E., Ruiz, J. D. J. L., Coppens d'Eeckenbrugge, G., et al. (2014). Multiple lines of evidence for the origin of domesticated chili pepper, *Capsicum annuum*, in Mexico. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6165–6170. doi: 10.1073/pnas.1308933111
- Kwak, M., and 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. doi: 10.1007/s00122-008-0955-4
- Kwak, M., Kami, J., and Gepts, P. (2009). The putative Mesoamerican domestication center of *Phaseolus vulgaris* is located in the Lerma-Santiago Basin of Mexico. *Crop Sci.* 49, 554–563. doi: 10.2135/cropsci2008.07.0421
- Lam, H. M., Xu, X., Liu, X., Chen, W., Yang, G., Wong, F. L., et al. (2010). Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nat. Genet.* 42, 1053–1059. doi: 10.1038/ng.715
- Lester, R. N. (1989). Evolution under domestication involving disturbance of genetic balance. *Euphytica* 44, 125–132. doi: 10.1007/BF00022606
- Lewontin, R. C., and Birch, L. C. (1966). Hybridization as a source of variation for adaptation to new environments. *Evolution* 20, 315–336. doi: 10.2307/2406633
- Lewontin, R. C., and Krakauer, J. (1973). Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphism. *Genetics* 74, 175–195.
- Liljegren, S., Ditta, G., Eshed, Y., Savidge, B., Bowmanand, J., and Yanofsky, M. (2000). Control of fruit dehiscence in *Arabidopsis* by the SHATTERPROOF MADS-box genes. *Nature* 404, 766–769. doi: 10.1038/35008089
- Lin, Z., Li, X., Shannon, L. M., Yeh, C. T., Wang, M. L., Bai, G., et al. (2012). Parallel domestication of the *Shattering1* genes in cereals. *Nat. Genet.* 44, 720–724. doi: 10.1038/ng.2281
- Lioi, L. (1989). Geographical variation of phaseolin patterns in an old world collection of *Phaseolus vulgaris*. *Seed Sci. Technol.* 17, 317–324.
- Lioi, L., Lotti, C., and Galasso, I. (1998). Isozyme diversity, RFLP of the rDNA and phylogenetic affinities among cultivated lima beans, *Phaseolus lunatus* (Fabaceae). *Plant Syst. Evol.* 213, 153–164. doi: 10.1007/BF00985196
- Lioi, L., and Piergiovanni, A. R. (2013). “European common bean,” in *Genetic and Genomic Resources of Grain Legume Improvement*, eds M. Singh, H. D. Upadhyaya, and I. S. Bisht (London: Elsevier Inc.), 11–40. doi: 10.1016/B978-0-12-397935-3.00002-5
- Lippold, S., Knapp, M., Kuznetsova, T., Leonard, J. A., Benecke, N., Ludwig, A., et al. (2011). Discovery of lost diversity of paternal horse lineages using ancient DNA. *Nat. Commun.* 2, 450. doi: 10.1038/ncomms1447
- Logozzo, G., Donnoli, R., Macaluso, L., Papa, R., Knüpfner, H., and Spagnoletti Zeuli, P. S. (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. *Genet. Res. Crop Evol.* 54, 1763–1779. doi: 10.1007/s10722-006-9185-2
- Londo, J. P., Chiang, Y. C., Hung, K. H., Chiang, T. Y., and Schaal, B. A. (2006). Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 9578–9583. doi: 10.1073/pnas.0603152103
- Lu, J., Tang, T., Tang, H., Huang, J., Shi, S., and Wu, C. I. (2006). The accumulation of deleterious mutations in rice genomes: a hypothesis on the cost of domestication. *Trends Genet.* 22, 126–131. doi: 10.1016/j.tig.2006.01.004
- Mamidi, S., Rossi, M., Annam, D., Moghaddam, S., Lee, R., Papa, R., et al. (2011). Investigation of the domestication of common bean (*Phaseolus vulgaris*) using multilocus sequence data. *Funct. Plant Biol.* 38, 953–967. doi: 10.1071/FP11124
- Mamidi, S., Rossi, M., Moghaddam, S. M., Annam, D., Lee, R., Papa, R., et al. (2013). Demographic factors shaped diversity in the two gene pools of wild common bean *Phaseolus vulgaris* L. *Heredity* 110, 267–276. doi: 10.1038/hdy.2012.82
- Manshardt, R. M., and Waines, J. G. (1983). Isozyme variation and the origin of domesticated tepary beans (*Phaseolus acutifolius* Gray). *Annu. Rep. Bean Improv. Coop.* 26, 18–19.
- Maquet, A., Zoro Bi, I., Delvaux, M., Wathelet, B., and Baudoin, J. P. (1997). Genetic structure of a lima bean base collection using allozyme markers. *Theor. Appl. Genet.* 95, 980–991. doi: 10.1007/s001220050651
- Márquez, A., Verma, S. P., Anguita, F., Oyarzun, R., and Brandle, J. L. (1999). Tectonics and volcanism of sierra chichinautzin: extension at the front of the central trans-mexican volcanic belt. *J. Volcanol. Geotherm. Res.* 93, 125–150. doi: 10.1016/S0377-0273(99)00085-2
- Martin, G. B., and Adams, M. W. (1987a). Landraces of *Phaseolus vulgaris* (Fabaceae) in northern Malawi I. Regional variation. *Econ. Bot.* 41, 190–203. doi: 10.1007/BF02858965
- Martin, G. B., and Adams, M. W. (1987b). Landraces of *Phaseolus vulgaris* (Fabaceae) in northern Malawi II. Generation and maintenance of variability. *Econ. Bot.* 41, 204–215. doi: 10.1007/BF02858967
- Martínez-Castillo, J., Camacho-Pérez, L., Villanueva-Viramontes, S., Andueza-Noh, R. H., and Chacón-Sánchez, M. I. (2014). Genetic structure within the Mesoamerican gene pool of wild *Phaseolus lunatus* (Fabaceae) from Mexico as revealed by microsatellite markers: implications for conservation and the domestication of the species. *Am. J. Bot.* 101, 851–864. doi: 10.3732/ajb.1300412
- Matsuoka, Y., Vigouroux, Y., Goodman, M. M., Sanchez, G. J., Buckler, E., and Doebley, J. (2002). A single domestication for maize shown by multilocus microsatellite genotyping. *Proc. Natl. Acad. Sci. U.S.A.* 99, 6080–6084. doi: 10.1073/pnas.052125199
- McCouch, S. (2004). Diversifying selection in plant breeding. *PLoS Biol.* 2:e347. doi: 10.1371/journal.pbio.0020347
- Mendel, G. (1866). Versuche über Pflanzenhybriden. *Verh. Naturf. Verein. Brünn* 4, 1–47.
- Messina, V. (2014). Nutritional and health benefits of dried beans. *Am. J. Clin. Nutr.* 100, 437S–442S. doi: 10.3945/ajcn.113.071472
- Meyer, R. S., DuVal, A. E., and Jensen, H. R. (2012). Patterns and processes in crop domestication: an historical review and quantitative analysis of 203 global food crops. *New Phytol.* 196, 29–48. doi: 10.1111/j.1469-8137.2012.04253.x
- Mizoguchi, T., Wright, L., Fujiwara, S., Cremer, F., Lee, K., Onouchi, H., et al. (2005). Distinct roles of GIGANTEA in promoting flowering and regulating circadian rhythms in *Arabidopsis*. *Plant Cell* 17, 2255–2270. doi: 10.1105/tpc.105.033464
- Molina, J., Sikora, M., Garud, N., Flowers, J. M., Rubinstein, S., Reynolds, A., et al. (2011). Molecular evidence for a single evolutionary origin of domesticated rice. *Proc. Natl. Acad. Sci. U.S.A.* 108, 8351–8356. doi: 10.1073/pnas.1104686108
- Motta-Aldana, J., Serrano-Serrano, M. L., Torres, H. J., Villamizar, C. G., Debouck, D. G., and Chacón, M. I. (2010). Multiple origins of Lima bean landraces in the Americas: evidence from chloroplast and nuclear DNA polymorphisms. *Crop Sci.* 50, 1773–1787. doi: 10.2135/cropsci2009.12.0706
- Muñoz, L. C., Duque, M. C., Debouck, D. G., and Blair, M. W. (2006). Taxonomy of tepary bean and wild relatives as determined by amplified fragment length polymorphism (AFLP) markers. *Crop Sci.* 46, 1744–1754. doi: 10.2135/cropsci2005-12-0475
- Murgia, M. L., Attene, G., Rodriguez, M., Bitocchi, E., Bellucci, E., Fois, D., et al. (2017). A comprehensive phenotypic investigation of the ‘pod-shattering syndrome’ in common bean. *Front. Plant Sci.* 8:251. doi: 10.3389/fpls.2017.00251
- Nabhan, G. P., and Felger, R. S. (1978). Teparies in southwestern North America. *Econ. Bot.* 32, 3–19. doi: 10.1007/BF02906725
- Nanni, L., Bitocchi, E., Bellucci, E., Rossi, M., Rau, D., Attene, G., et al. (2011). Nucleotide diversity of a genomic sequence similar to SHATTERPROOF (*PvSHP1*) in domesticated and wild common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 123, 1341–1357. doi: 10.1007/s00122-011-1671-z
- Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89, 583–590.
- Nieto-Samaniego, A., Ferrari, L., Alaniz-Alvarez, S., Labarthe-Hernandez, G., and Rosas-Elguera, J. (1999). Variation of Cenozoic extension and volcanism across the southern Sierra Madre Occidental volcanic province, Mexico. *Geol. Soc. Am. Bull.* 111, 347–363. doi: 10.1130/0016-7606
- Nowosielski, J., Podyma, W., and Nowosielska, D. (2002). Molecular research on the genetic diversity of Polish varieties and landraces of *Phaseolus coccineus* L. and *Phaseolus vulgaris* L. using the RAPD and AFLP methods. *Cell. Mol. Biol. Lett.* 7, 753–762.
- Okii, D., Tukamuhawa, P., Kami, J., Namayanja, A., Paparu, P., Ugen, M., et al. (2014). The genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) germplasm in Uganda. *Afr. J. Biotechnol.* 13, 2935–2949. doi: 10.5897/AJB2014.13916

- Olson, M. V. (1999). When less is more: gene loss as an engine of evolutionary change. *Am. J. Hum. Genet.* 64, 18–23. doi: 10.1086/302219
- Osakabe, Y., Arinaga, N., Umezawa, T., Katsura, S., Nagamachi, K., Tanaka, H., et al. (2013). Osmotic stress responses and plant growth controlled by potassium transporters in *Arabidopsis*. *Plant Cell* 25, 609–624. doi: 10.1105/tpc.112.105700
- Papa, R., Acosta, J., Delgado-Salinas, A., and Gepts, P. (2005). A genome-wide analysis of differentiation between wild and domesticated *Phaseolus vulgaris* from Mesoamerica. *Theor. Appl. Genet.* 111, 1147–1158. doi: 10.1007/s00122-005-0045-9
- Papa, R., Bellucci, E., Rossi, M., Leonardi, S., Rau, D., Gepts, P., et al. (2007). Tagging the signatures of domestication in common bean (*Phaseolus vulgaris*) by means of pooled DNA samples. *Ann. Bot.* 100, 1039–1051. doi: 10.1093/aob/mcm151
- 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. *Theor. Appl. Genet.* 106, 239–250. doi: 10.1007/s00122-002-1085-z
- Papa, R., Nanni, L., Sicard, D., Rau, D., and Attene, G. (2006). “The evolution of genetic diversity in *Phaseolus vulgaris* L.” in *New Approaches to the Origins, Evolution and Conservation of Crops. Darwin’s Harvest*, eds T. J. Motley, N. Zerega, and H. Cross (New York, NY: Columbia University Press), 121–142.
- Plunket, P., and Uruñuela, G. (2012). Where east meets west: the formative in Mexico’s central highlands. *J. Archaeol. Res.* 20, 1–51. doi: 10.1007/s10814-011-9051-4
- Purugganan, M. D., and Fuller, D. (2009). The nature of selection during plant domestication. *Nature* 457, 843–848. doi: 10.1038/nature07895
- Raggi, L., Tissi, C., Mazzucato, A., and Negri, V. (2014). Molecular polymorphism related to flowering trait variation in a *Phaseolus vulgaris* L. collection. *Plant Sci.* 215, 180–189. doi: 10.1016/j.plantsci.2013.11.001
- Ramírez-Villegas, J., Khoury, C., Jarvis, A., Debouck, D. G., and Guarino, L. (2010). A gap analysis methodology for collecting crop gene pools: a case study with *Phaseolus* beans. *PLoS ONE* 5:e13497. doi: 10.1371/journal.pone.0013497
- Rana, J. C., Sharma, T. R., Tyagi, R. K., Chahota, R. K., Gautam, N. K., Singh, M., et al. (2015). Characterisation of 4274 accessions of common bean (*Phaseolus vulgaris* L.) germplasm conserved in the Indian gene bank for phenological, morphological and agricultural traits. *Euphytica* 205, 441–457. doi: 10.1007/s10681-015-1406-3
- Rendón-Anaya, M., Montero-Vargas, J. M., Saburido-Álvarez, S., Vlasova, A., Capella-Gutiérrez, S., Ordaz-Ortiz, J. J., et al. (2017). Genomic history of the origin and domestication of common bean unveils its closest sister species. *Genome Biol.* 18, 60. doi: 10.1186/s13059-017-1190-6
- Rodino, A. P., Lema, E. M., Marlene, P. B., Santalla, M., and De Ron, A. M. (2007). Assessment of runner bean (*Phaseolus coccineus* L.) germplasm for tolerance to low temperature during early seedling growth. *Euphytica* 155, 63–70. doi: 10.1007/s10681-006-9301-6
- Rodriguez, M., Rau, D., Angioi, S. A., Bellucci, E., Bitocchi, E., Nanni, L., et al. (2013). European *Phaseolus coccineus* L. landraces: population structure and adaptation, as revealed by cpSSRs and phenotypic analyses. *PLoS ONE* 8:e57337. doi: 10.1371/journal.pone.0057337
- Rodriguez, M., Rau, D., Bitocchi, E., Bellucci, E., Biagetti, E., Carboni, A., et al. (2016). Landscape genetics, adaptive diversity, and population structure in *Phaseolus vulgaris*. *New Phytol.* 209, 1781–1794. doi: 10.1111/nph.13713
- Rossi, M., Bitocchi, E., Bellucci, E., Nanni, L., Rau, D., Attene, G., et al. (2009). Linkage disequilibrium and population structure in wild and domesticated populations of *Phaseolus vulgaris* L. *Evol. Appl.* 2, 504–522. doi: 10.1111/j.1752-4571.2009.00082.x
- Rubiales, D., and Mikic, A. (2015). Introduction: legumes in sustainable agriculture. *Crit. Rev. Plant Sci.* 34, 2–3. doi: 10.1080/07352689.2014.897896
- Salunkhe, D. K., and Kadam, S. S. (1998). *Handbook of Vegetable Science and Technology: Production, Compostion, Storage, and Processing*. Boca Raton, FL: CRC press.
- Santalla, M., De Ron, A. M., and De La Fuente, M. (2010). Integration of genome and phenotypic scanning gives evidence of genetic structure in Mesoamerican common bean (*Phaseolus vulgaris* L.) landraces from the southwest of Europe. *Theor. Appl. Genet.* 120, 1635–1651. doi: 10.1007/s00122-010-1282-0
- Santalla, M., Monteagudo, A. B., Gonzalez, A. M., and De Ron, A. M. (2004). Agronomical and quality traits of runner bean germplasm and implications for breeding. *Euphytica* 135, 205–215. doi: 10.1023/B:EUPH.0000014912.07993.e7
- Santalla, M., Rodiño, A. P., and De Ron, A. M. (2002). Allozyme evidence supporting southwest Europe as a secondary center of genetic diversity for common bean. *Theor. Appl. Genet.* 104, 934–944. doi: 10.1007/s00122-001-0844-6
- Schinkel, C., and Gepts, P. (1988). Phaseolin diversity in the tepary bean *Phaseolus acutifolius* A. Gray. *Plant Breed.* 101, 292–301. doi: 10.1111/j.1439-0523.1988.tb00301.x
- Schinkel, C., and Gepts, P. (1989). Allozyme variability in the tepary bean, *Phaseolus acutifolius* A. Gray. *Plant Breed.* 102, 182–195. doi: 10.1111/j.1439-0523.1989.tb00336.x
- Schmit, V., and Debouck, D. G. (1991). Observations on the origin of *Phaseolus polyanthus* Greenman. *Econ. Bot.* 45, 345–364. doi: 10.1007/BF02887077
- Schmutz, J., McClean, P. E., Mamidi, S., Wu, G. A., Cannon, S. B., Grimwood, J., et al. (2014). A reference genome for common bean and genome-wide analysis of dual domestications. *Nat. Genet.* 46, 707–713. doi: 10.1038/ng.3008
- Serrano-Serrano, M. L., Andueza-Noh, R. H., Martínez-Castillo, J., Debouck, D. G., and Chacón, M. I. (2012). Evolution and domestication of Lima bean (*Phaseolus lunatus* L.) in Mexico: Evidence from ribosomal DNA. *Crop Sci.* 52, 1698–1712. doi: 10.2135/cropsci2011.12.0642
- Serrano-Serrano, M. L., Hernandez-Torres, J., Castillo-Villamizar, G., Debouck, D. G., and Chacón, M. I. (2010). Gene pools in wild Lima bean (*Phaseolus lunatus* L.) from the Americas: evidences for an Andean origin and past migrations. *Mol. Phylogenet. Evol.* 54, 76–87. doi: 10.1016/j.ympev.2009.08.028
- Shii, C. T., Rabakoarihanta, A., Mok, M. C., and Mok, D. W. S. (1982). Embryo development in reciprocal crosses of *Phaseolus vulgaris* and *Phaseolus coccineus*. *Theor. Appl. Genet.* 62, 59–64. doi: 10.1007/BF00276284
- 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 Breed.* 124, 464–472. doi: 10.1111/j.1439-0523.2005.01137.x
- Siebe, C. (2000). Age and archaeological implications of xitle volcano, southwestern basin of Mexico-city. *J. Volcanol. Geotherm. Res.* 104, 45–64. doi: 10.1016/S0377-0273(00)00199-2
- Siebe, C., Rodriguez-Lara, V., Schaaf, P., and Abrams, M. (2004). Radiocarbon ages of Holocene Pelado, Guespalapa, and Chichinautzin scoria cones, south of Mexico City: implications for archaeology and future hazards. *Bull. Volcanol.* 66, 203–225. doi: 10.1007/s00445-003-0304-z
- Sofi, P. A., Rana, J. C., and Bhat, N. A. (2014). Pattern of variation in common bean (*Phaseolus vulgaris* L.) genetic resources of Jammu and Kashmir. *J. Food Legumes* 27, 197–201.
- Spataro, G., Tiranti, B., Arcaleni, P., Bellucci, E., Attene, G., Papa, R., et al. (2011). Genetic diversity and structure of a worldwide collection of *Phaseolus coccineus* L. *Theor. Appl. Genet.* 122, 1281–1291. doi: 10.1007/s00122-011-1530-y
- Swanson-Wagner, R., Briskine, R., Schaefer, R., Hufford, M. B., Ross-Ibarra, J., Myers, C. L., et al. (2012). Reshaping of the maize transcriptome by domestication. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11878–11883. doi: 10.1073/pnas.1201961109
- Tanksley, S. D., and McCouch, S. R. (1997). Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277, 1063–1066. doi: 10.1126/science.277.5329.1063
- Tarrago, M. N. (1980). “El proceso de agriculturización en el Noroeste Argentino, zona Valliserrana,” in *Actas del V Congreso Nacional Arqueología Argentina*, Vol. 1, (San Juan: Universidad de San Juan), 181–217.
- Toro, O., Tohme, J., and Debouck, D. G. (1990). *Wild Bean (Phaseolus vulgaris L.): Description and Distribution*. Cali: Centro Internacional de Agricultura Tropical.
- van Heerwaarden, J., Doebley, J., Briggs, W. H., Glaubitz, J. C., Goodman, M. M., Sanchez-Gonzalez, J. J., et al. (2011). Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proc. Natl. Acad. Sci. U.S.A.* 108, 1088–1092. doi: 10.1073/pnas.1013011108
- Vaz Patto, M. C., Amarowicz, R., Aryee, A. N., Boye, J. I., Chung, H. J., Martín-Cabrejas, M. A., et al. (2015). Achievements and challenges in improving the nutritional quality of food legumes. *Crit. Rev. Plant Sci.* 34, 105–143. doi: 10.1080/07352689.2014.897907

- Vitte, C., Ishii, T., Lamy, F., Brar, D., and Panaud, O. (2004). Genomic paleontology provides evidence for two distinct origins of Asian rice (*Oryza sativa* L.). *Mol. Genet. Genom.* 272, 504–511. doi: 10.1007/s00438-004-1069-6
- Vlasova, A., Capella-Gutiérrez, S., Rendón-Anaya, M., Hernández-Oñate, M., Minoche, A. E., Erb, I., et al. (2016). Genome and transcriptome analysis of the Mesoamerican common bean and the role of gene duplications in establishing tissue and temporal specialization of genes. *Genome Biol.* 17, 32. doi: 10.1186/s13059-016-0883-6
- vonHoldt, B. M., Pollinger, J. P., Lohmueller, K. E., Han, E., Parker, H. G., Quignon, P., et al. (2010). Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* 464, 898–902. doi: 10.1038/nature08837
- Vrebalov, J., Panb, I. L., Arroyo, A. J., McQuinn, R., Chung, M., Poole, M., et al. (2009). Fleshy fruit expansion and ripening are regulated by the tomato *SHATTERPROOF* gene *TAGL1*. *Plant Cell* 21, 3041–3062. doi: 10.1105/tpc.109.066936
- Wall, J. R. (1970). Experimental introgression in the genus *Phaseolus*. 1. Effect of mating systems on interspecific gene flow. *Evolution* 24, 356–366. doi: 10.2307/2406810
- Wang, I. J., Glor, R. E., and Losos, J. B. (2013). Quantifying the roles of ecology and geography in spatial genetic divergence. *Ecol. Lett.* 16, 175–182. doi: 10.1111/ele.12025
- Westphal, E. (ed.) (1974). “*Phaseolus*,” in *Pulses in Ethiopia, their Taxonomy and Agricultural Significance*, (Wageningen: Centre for Agricultural Publishing and Documentation, (PUDOC)), 129–176.
- Wortmann, C. S., Kirkby, R. A., Eledu, C. A., and Allen, D. J. (1998). *Atlas of Common Bean (Phaseolus vulgaris L.) Production in Africa*. Cali: Centro Internacional de Agricultura Tropical (CIAT), 133.
- Xia, Q., Guo, Y., Zhang, Z., Li, D., Xuan, Z., Li, Z., et al. (2009). Complete resequencing of 40 genomes reveals domestication events and genes in silkworm (*Bombyx*). *Science* 326, 433–436. doi: 10.1126/science.1176620
- Xu, X., Liu, X., Ge, S., Jensen, J. D., Hu, F., Li, X., et al. (2012). Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat. Biotechnol.* 30, 105–111. doi: 10.1038/nbt.2050
- Young, K. R., Ulloa Ulloa, C., and Lutey, J. L. (2002). Plant evolution and endemism in Andean South America: an introduction. *Bot. Rev.* 68, 4–21. doi: 10.1663/0006-8101(2002)068[0004:PEAEIA]2.0.CO;2
- Zhang, X., Blair, M. W., and Wang, S. (2008). Genetic diversity of Chinese common bean (*Phaseolus vulgaris* L.) landraces assessed with simple sequence repeats markers. *Theor. Appl. Genet.* 117, 629–640. doi: 10.1007/s00122-008-0807-2
- Zizumbo-Villarreal, D., and Colunga-GarcíaMarín, P. (2010). Origin of agriculture and plant domestication in West Mesoamerica. *Genet. Res. Crop Evol.* 57, 813–825. doi: 10.1007/s10722-009-9521-4
- Zizumbo-Villarreal, D., Papa, R., Hufford, M., Repinski, S., and Gepts, P. (2009). Identification of a new wild population of *Phaseolus vulgaris* L. in western Jalisco, Mexico, near the Mesoamerican domestication center of common bean. *BIC* 52, 24–25.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Bitocchi, Rau, Bellucci, Rodriguez, Murgia, Gioia, Santo, Nanni, Attene and Papa. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.