Chloroplast microsatellite diversity in *Phaseolus vulgaris*

F. Desiderio¹, E. Bitocchi¹, E. Bellucci¹, D. Rau², M. Rodríguez², G. Attene²,³, R. Papa¹,⁴ and L. Nanni¹*  

¹ Dipartimento di Scienze Agrarie, Alimentari ed Ambientali, Università Politecnica delle Marche, Ancona, Italy 
² Dipartimento di Agraria, Università degli Studi di Sassari, Sassari, Italy 
³ Centro per la Conservazione e la Valorizzazione della Biodiversità Vegetale, Università degli Studi di Sassari, Surigheddu, Alghero, Italy 
⁴ Cereal Research Centre, Consiglio per la Ricerca e Sperimentazione in Agricoltura, Foggia, Italy  

*Correspondence: L. Nanni, Dipartimento di Scienze Agrarie, Alimentari ed Ambientali, Università Politecnica delle Marche, Via Brecce Bianche, 60131 Ancona, Italy. e-mail: l.nanni@unipolm.it

Evolutionary studies that are aimed at defining the processes behind the present level and organization of crop genetic diversity represent the fundamental bases for biodiversity conservation and use. A Mesoamerican origin of the common bean *Phaseolus vulgaris* was recently suggested through analysis of nucleotide polymorphism at the nuclear level. Here, we have used chloroplast microsatellites to investigate the origin of the common bean, on the basis of the specific characteristics of these markers (no recombination, haploid genome, uniparental inheritance), to validate these recent findings. Indeed, comparisons of the results obtained through analysis of nuclear and cytoplasmic DNA should allow the resolution of some of the contrasting information available on the evolutionary processes. The main outcomes of the present study are: (i) confirmation at the chloroplast level of the results obtained through nuclear data, further supporting the Mesoamerican origin of *P. vulgaris*, with central Mexico representing the cradle of its diversity; (ii) identification of a putative ancestral plastidial genome, which is characteristic of a group of accessions distributed from central Mexico to Peru, but which have not been highlighted beforehand through analyses at the nuclear level. Finally, the present study suggests that when a single species is analyzed, there is the need to take into account the complexity of the relationships between *P. vulgaris* and its closely related and partially intercrossable species *P. coccineus* and *P. dumosus*. Thus, the present study stresses the importance for the investigation of the speciation processes of these taxa through comparisons of both plastidial and nuclear variability. This knowledge will be fundamental not only from an evolutionary point of view, but also to put *P. coccineus* and *P. dumosus* germplasm to better use as a source of useful diversity for *P. vulgaris* breeding.

Keywords: Phaseolus, crop evolution, cpSSR, recombination, population structure, speciation, introgression

INTRODUCTION

The wild forms of the common bean *Phaseolus vulgaris* grow across a wide geographic area of the Americas, from northern Mexico to northwestern Argentina (Toro et al., 1990). Morphological, biochemical, and molecular data have indicated that the wild populations from Mexico, Central America, and Colombia differ from those of southern Peru, Bolivia, and Argentina (Gepts et al., 1986; Delgado-Salinas et al., 1988; Koenig and Gepts, 1989; Gepts and Debouck, 1991; Becerra-Velasquez and Gepts, 1994; Papa and Gepts, 2003; Angioi et al., 2009a; Kwak and Gepts, 2009; Rossi et al., 2009). Indeed, these two groups represent two geographically distinct and isolated gene pools (Mesoamerica and Andes, respectively) that were already present before domestication of the common bean (for reviews, see Papa et al., 2006; Bitocchi et al., 2012, 2013). This complex scenario is further characterized by the presence within the wild forms of a third gene pool that is characteristic of a restricted area in northern Peru and Ecuador (Debouck et al., 1993). Along with accessions from the two main gene pools, wild populations collected in this restricted area have been analyzed according to a portion of the gene encoding for the seed-storage protein phaseolin (Kami et al., 1995). This study showed that the “Inca” phaseolin type I is not present in Central and South America. Moreover, this phaseolin appears to be ancestral to the other phaseolin sequences of *P. vulgaris*, suggesting that the northern Peru and Ecuador populations were those from which the common bean originated and subsequently spread into Central and South America (Kami et al., 1995). This hypothesis was the most credited until the study of Bitocchi et al. (2012) that analyzed the genetic diversity at five nuclear gene fragments across a wide sample of wild *P. vulgaris* accessions, where they showed that the wild forms of *P. vulgaris* originated in Mesoamerica, and most likely in central Mexico. This study also indicated that both the Andean and the northern Peru and Ecuador gene pools originated through different migration events from central Mexico. This conclusion was suggested by the evidence of a bottleneck that occurred in the Andes prior to domestication (Rossi et al., 2009; Nanni et al., 2011; Bitocchi et al., 2012) and to the presence of high genetic structure in Mesoamerica (Bitocchi et al., 2012), with the different genetic groups identified having diverse relationships with the wild populations from northern Peru and Ecuador and from the Andes.

Chloroplast microsatellite (cpSSR) markers are widely used in population genetics and evolutionary studies of plants (for review, see Provan et al., 2001). Due to their specific characteristics, which
include a haploid and non-recombinant genome and uniparental inheritance, they have become very useful tools to investigate different evolutionary processes. These include, e.g., historical bottlenecks, founder effects, identification of progenitors of cultivated species, and the role of introgression in crop evolution (for review, see Provan et al., 2001).

In the present study, we used a set of cpSSRs to analyze a wide sample of wild *P. vulgaris* accessions from the Americas. These cpSSRs have been demonstrated to be very useful to study the diversity and evolution of several legume species, and in particular of *P. vulgaris* and *P. coccineus* (Angioi et al., 2009a,b, 2010). The main aim was to investigate the origin of the common bean and to compare the results with those obtained by analyses based on nuclear nucleotide diversity (Bitocchi et al., 2012). Indeed, at the nuclear level, recombination might have affected the data obtained, although to reduce this problem, fragments of a few hundreds of base pairs were used. Thus, the comparison and combination of nuclear and plastidial polymorphism analyses should give complementary insights into the evolutionary history of the common bean, especially considering that such analyses can often provide contrasting information on evolutionary processes (Birky, 1988; McCauley, 1995; Ennos et al., 1999; Provan et al., 1999; Weising and Gardner, 1999; Ishii et al., 2001; Lira et al., 2003; Ueno et al., 2005).

Finally, cpSSR genotyping of a smaller set of *P. coccineus* accessions was carried out, with the aim being to gain information about the evolutionary relationship between *P. coccineus* and *P. vulgaris*.

**MATERIALS AND METHODS**

**PLANT MATERIALS**

A total of 109 wild accessions of *P. vulgaris* were analyzed in the present study. These materials encompassed the entire geographical distribution of this species, from northern Mexico to northwestern Argentina, and included seven wild accessions from northern Peru and Ecuador that are characterized by the ancestral phaseolin type I (Debouck et al., 1993; Kami et al., 1995). The geographical distribution of these common bean accessions is shown in Figure 1. Ten wild accessions of *P. coccineus* were also included. Each accession is represented by an individual plant genotype. A complete list of the accessions studied, along with their “passport” information, is given in Table A1 in Appendix.

The seeds were provided by the United States Department of Agriculture (USDA) Western Regional Plant Introduction Station in the USA, the International Center of Tropical Agriculture (CIAT) in Colombia, and the Laboratory of Plant Genetics (D3A) at the Polytechnic University of Marche (UNIVPM) in Italy. Most of these accessions had already been characterized using different types of molecular markers, such as amplified fragment length polymorphism (AFLP; Rossi et al., 2009) and nucleotide data (Nanni et al., 2011; Bitocchi et al., 2012). Moreover a small subset of accessions (15 wild *P. vulgaris*, eight wild *P. coccineus*) were analyzed previously by Angioi et al. (2009a) with the same set of cpSSRs.

**PCR AND cpSSR GENOTYPING**

Genomic DNA was extracted from each accession from young leaf tissue of a single, greenhouse-grown plant, using the miniprep extraction method (Doyle and Doyle, 1987). A total of 17 cpSSRs derived from the literature (Weising and Gardner, 1999; Chung and Staub, 2003; Angioi et al., 2009a) were used for the genetic characterization of the whole sample. One of the two SSR primers was end-labeled with a phosphoramidite fluorescent dyes, 6-FAM or HEX. A list of the cpSSRs used in this study is given Table A2 in Appendix. The amplifications were conducted using a Perkin-Elmer 9700 Thermal Cycler (PE Applied Biosystems) in a total volume of 25 µl, which contained 25 ng template DNA, 10 pmol of each primer, 200 µM dNTPs, 1 × Taq polymerase buffer, 2.5 mM MgCl₂ and 1 U Taq polymerase (Promega). The PCR conditions were as reported in Table A2 in Appendix. Multiplex PCRs were performed (including two primer pairs that were differently end-labeled, with amplification of SRRs of different sizes under the same amplification conditions). SSR genotyping was carried out using the ABI PRISM 3100-Avant Genetic Analyzer with the GeneScan 7.0 analysis software (PE Applied Biosystems).

**GENETIC DIVERSITY ANALYSIS**

The percentage of polymorphic loci, the average number of observed alleles per locus (Na), the effective number of alleles per locus (Ne; Kimura and Crow, 1964), the number of private alleles (Np), and the expected heterozygosity (He; Nei, 1978) estimates based on allele frequencies, were computed using the Arlequin software, version 3.5 (Excoffier and Lischer, 2010). The whole sample, and the following partitions of the accessions were considered for these analyses: *P. coccineus*; *P. vulgaris*; and within the common bean sample according to the gene pool, the Andean wild (AW), Mesoamerican wild (MW), and northern Peru and Ecuador (PhI) populations.
The differences between the AW and MW populations for the genetic diversity estimates (Ne and He) were tested using Wilcoxon signed-ranks non-parametric test for two groups, arranged for paired observations (i.e., one pair of estimates for each locus; Wilcoxon, 1945; Sokal and Rohlf, 1995).

An ad hoc statistic ($\Delta H$) was used to compare the diversity between the two main gene pools (AW, MW); this estimate measures the loss of diversity of one population compared to another, and it was originally proposed by Vigouroux et al. (2002): $\Delta H = 1 - (H_{POP1}/H_{POP2})$, where POP1 refers to the population that shows the lower level of genetic diversity (He) compared to the other population (POP2).

PRINCIPAL COMPONENT ANALYSIS

Using the JMP software, version 8 (SAS Institute, Inc., 2008), principal component analysis (PCA) was performed from allele frequencies. The same analysis was carried out also to investigate the genetic relationships among the *P. vulgaris* accessions.

POPULATION STRUCTURE ANALYSIS

A Bayesian model-based approach that was implemented in the Bayesian analysis of population structure (BAPS) software, version 5.3 (Corander et al., 2003), was used to infer the hidden genetic population structure of the whole sample (109 *P. vulgaris* and 10 *P. coccineus* accessions), and thus to assign the genotypes into genetically structured groups/populations (K). A spatial genetic mixture analysis was conducted (Corander et al., 2008). This method uses a Markov chain Monte Carlo simulation approach to group samples into variable user-defined numbers (K) of clusters. The best partition of populations into K clusters is identified as the one with the highest marginal log-likelihood. We carried out 10 repetitions of the algorithm for each K ranging between 2 and 20.

The genetic diversity statistics described above were also computed for the genetic groups highlighted by the BAPS analysis (hereafter referred to as clusters). The differences between the clusters identified according to the genetic diversity estimates (Ne, He) were tested using the Wilcoxon signed-ranks non-parametric test for two groups, arranged for paired observations (Wilcoxon, 1945; Sokal and Rohlf, 1995), and the Bonferroni correction for multiple comparisons.

DIVERGENCE BETWEEN POPULATIONS

The divergence among the *P. coccineus* and *P. vulgaris* populations defined *a priori* according to the gene pools (AW, MW, PHI) were estimated as $F_{ST}$ (Weir and Cockerham, 1984), $D$ (Jost, 2008), and $R_{ST}$ (Slatkin, 1995). In contrast to $F_{ST}$ and $D$, $R_{ST}$ contains information not only about the frequency with which particular alleles occur, but also on the evolutionary distance between them, inasmuch as it is measured as the expected squared difference in repeat numbers between alleles. For this reason, it is intended to take advantage of this additional information to provide greater insight into the patterns of relationships among populations (for review, see Holsinger and Weir, 2009). These correspond to the infinite allele and the step-wise mutation models. The significance of the estimates was obtained through permutation tests, using 10,000 permutations. The same divergence estimates were also computed for clusters identified by BAPS analysis. The Arlequin software, version 3.5 (Excoffier and Lischer, 2010), was used.

COMPARISON OF RESULTS BASED ON cpSSR DATA WITH THOSE OBTAINED USING NUCLEOTIDE DATA

The sequences of five gene regions (from 500 to 900 bp) for 71 accessions were available from Bitocchi et al. (2012). These five gene fragments include four legume anchor (Leg) markers, developed by Hougaard et al. (2008), and one gene fragment, *PvSHP1*, developed by Nanni et al. (2011); *PvSHPi1* is a homolog of the SHATTERPROOF (SHP1) gene, which is involved in the control of fruit splitting in *Arabidopsis thaliana*. These data allowed a comparison of the data from the population structure analyses carried out using cpSSRs and nuclear sequences. Thus, for the 71 accessions shared between this study and the one of Bitocchi et al. (2012), a population structure analysis was carried out using both the cpSSRs and the nucleotide data. For the nucleotide data, the procedures were as described in Bitocchi et al. (2012), while for the cpSSRs, the procedures were the same as reported in the above section.

To compare the geographical distributions of the clusters identified through the cpSSR and nucleotide data, spatial interpolation of membership coefficients ($q$) was performed according to the kriging method, with each of the clusters identified by population structure analysis, which was implemented in the R packages spatial (http://www.r-project.org/). In the case of the cpSSRs, due to the non-recombinant nature of these markers, which does not allow admixture, the membership coefficients were represented by one or zero (i.e., membership or non-membership to one cluster); thus, the interpolation for plastidial data represents an approximation.

The association between the results obtained by the BAPS analyses carried out with the cpSSR and nucleotide data was tested by analysis of contingency tables with the likelihood ratio chi-squared ($\chi^2$) test, which was performed using the JMP 8.0 software (SAS Institute, Inc., 2008).

RESULTS

Each of the primer pairs produced a single and clear amplification, and all of the 17 loci studied were polymorphic considering the whole analyzed sample. The size of the amplification products ranged from 79 bp (ccmp3) to 378 bp (ccSSR19). Overall, the number of alleles per locus (Na) ranged from two (cp2) to 12 (ccSSR20); in parallel the same two markers showed the lowest and the highest genetic diversity, He = 0.13 and He = 0.85, respectively (Table A3 in Appendix).

Considering the *P. coccineus* sample, six out of the 17 loci were monomorphic. For the polymorphic loci, Na ranged from two (cp2, ccSSR2, ccSSR4, ccSSR12, and ccSSR16) to six (ccSSR20). One locus (cp2) was monomorphic in the *P. vulgaris* sample. For the remaining 16 loci, Na ranged from two (cp3 and ccSSR12) to 11 (ccSSR20). The highest level of genetic diversity was detected for the ccSSR20 locus, as an He of 0.84 for both *P. vulgaris* and *P. coccineus* (Table A3 in Appendix).

GENETIC DIVERSITY ANALYSIS

Genetic diversity estimates were computed considering the whole sample and the following major subdivisions: different species (*P. vulgaris*, *P. coccineus*) and within the *P. vulgaris* Andean (AW),
Table 1 | Genetic diversity estimates computed for all of the 17 cpSSR loci considering the whole sample, the *P. vulgaris* and *P. coccineus* samples, and the three *P. vulgaris* populations defined according to the gene pools.

| Accession | N  | % polymorphic loci | Na  | Ne  | Np (freq. ≥ 0.05) | He  |
|-----------|----|---------------------|-----|-----|-------------------|-----|
| All       | 119| 100                 | 5.1 | 2.6 | na                | 0.54|
| *P. vulgaris* | 109| 94.1                | 4.4 | 2.5 | 45                | 0.51|
| *P. coccineus* | 10 | 64.7                | 2.4 | 1.8 | 12                | 0.36|
| *P. vulgaris* populations | | | | | | |
| MW       | 55 | 88.2               | 3.9 | 2.5 | 7                 | 0.54|
| AW       | 47 | 82.4               | 3.2 | 1.9 | 4                 | 0.40|
| Phi      | 7  | 82.4               | 2.5 | 2.2 | 3                 | 0.49|

N, sample size; Na, mean number of observed alleles per locus; Ne, mean effective number of alleles per locus; Np, number of private alleles; Np (freq. ≥ 0.05), number of private alleles with frequency higher than 0.05; He, expected heterozygosity; MW, Mesoamerican wild; AW, Andean wild; Phi, northern Peru and Ecuador; na, not applicable.

Mesoamerican (MW), and northern Peru and Ecuador accession (Phi) populations.

As showed in Table 1, the common bean was characterized by a higher level of genetic diversity (Na, Ne, Np, and He) than *P. coccineus*. However, the large difference between the size of the two samples suggests caution in the consideration of these estimates.

Among the three *P. vulgaris* populations, the MW accessions showed the highest genetic diversity for all of the parameters (Table 1). In particular, considering the populations that represent the two major common bean gene pools (Mesoamerican and Andean), the MW showed a higher level of genetic diversity (Ne = 2.5 and He = 0.54) compared to the AW accessions (Ne = 1.9 and He = 0.40; Table 1). This difference was significant for both the genetic diversity estimates Ne and He (P < 0.02; Wilcoxon signed-ranks non-parametric test for two groups, arranged for paired observations). There was a 26% reduction in genetic diversity (ΔH) of the AW population compared to the MW population.

**PRINCIPAL COMPONENT ANALYSIS**

The relationships among all of the individuals considered, including both the *P. vulgaris* and *P. coccineus* accessions, were investigated by PCA (Figure 2). The first (PC1) and second (PC2) principal components explain 43.03 and 26.82%, respectively. Three main groups were identified by this analysis, one including eight wild *P. coccineus* accessions, one including all of the seven Phi, two WA, and 39 WM accessions and one *P. coccineus* accession, and the remaining 45 WA and 16 WM accessions, and even if more distant, one *P. coccineus* accessions.

Principal component analysis was also performed to investigate the genetic relationships among the *P. vulgaris* accessions (Figure 3). The first (PC1) and second (PC2) principal components explain 45.73 and 23.65%, respectively. This analysis identified two major groups, as A and B (Figure 3). The majority of the MW accessions (73%; including five of the six Colombian accessions) belonged to group A, along with three AW accessions from northern Argentina (Salta and Tucumán Provinces) and all of the seven Phi accessions. Group B included almost all of the AW accessions (94%) and 15 MW accessions, 14 of which were from central Mexico, and only one from Colombia.

**POPULATION STRUCTURE**

The population structure analysis identified four different clusters (C1, C2, C3, C4) as the best partition of the whole sample (all of the 10 best marginal log-likelihood values were for K = 4, with the highest of −1,996.54; Table 2). Cluster C1 was characterized by almost all of the AW accessions (98%) and 13 MW accessions from Central Mexico. Cluster C2 included 21 MW and three Phi accessions, along with two *P. coccineus* genotypes. There were accessions from all of the three common bean populations in cluster C3 (4, 1, 21 for the Phi, AW, MW populations, respectively), while cluster C4 was exclusive to the remaining eight *P. coccineus* accessions. The geographical distribution of the *P. vulgaris* accessions based on the BAPS cluster membership is showed in Figure 4.

The genetic diversity estimates for the BAPS clusters are showed in Table 3. The three clusters characteristic of *P. vulgaris* accessions (C1, C2, C3) showed similar levels of genetic diversity (Ne = 2.0, 2.1, 1.8, and He = 0.42, 0.45, 0.36, for C1, C2, C3, respectively). Cluster C4 showed the lowest Ne (1.6) and He (0.29) estimates. However, there were no significant differences in the levels of genetic diversity between these four clusters (Wilcoxon signed-ranks non-parametric tests, after Bonferroni correction).

**DIVERGENCE BETWEEN POPULATIONS**

The genetic divergence between the *P. vulgaris* populations (MW, AW, Phi) and the *P. coccineus* accessions was estimated as $F_{ST}$, $D$, and $R_{ST}$. The $F_{ST}$ and $D$ estimates were very similar, as expected for populations that have a very low number of unique alleles (Whitlock, 2011), and thus only the $F_{ST}$ data are shown. The lowest, and non-significant, differentiation was between the Phi and MW populations ($F_{ST} = 0.08$; $R_{ST} = 0.12$; both non-significant; Table 4). Considering the comparisons among the *P. vulgaris* populations, the divergence between AW and Phi ($F_{ST} = 0.21$; $R_{ST} = 0.70$; both significant $P ≤ 0.001$) was greater than that between AW and MW ($F_{ST} = 0.13$; $R_{ST} = 0.24$; both significant $P ≤ 0.01$). The highest values of $F_{ST}$ were those in the comparisons with the *P. coccineus* population; however, the MW population showed the lowest levels of differentiation with *P. coccineus* ($F_{ST} = 0.33$; $P ≤ 0.001$) compared to the other *P. vulgaris* populations [$F_{ST}(\text{Phi}-\text{P. coccineus}) = 0.38$, $P ≤ 0.001$; $F_{ST}(\text{AW}-\text{P. coccineus}) = 0.49$, $P ≤ 0.001$; Table 4]. The $R_{ST}$ showed a similar trend, with the MW population being less differentiated than *P. coccineus* ($R_{ST} = 0.58$, $P ≤ 0.001$), and Phi [$R_{ST}(\text{Phi}-\text{P. coccineus}) = 0.60$, $P ≤ 0.001$], and AW [$R_{ST}(\text{AW}-\text{P. coccineus}) = 0.78$, $P ≤ 0.001$; Table 4].

The same divergence estimates were computed considering the four genetic clusters (C1, C2, C3, C4) identified by the BAPS analysis (Table 5). All of the divergence estimates (for both $F_{ST}$ and $R_{ST}$) were significantly different from zero ($P ≤ 0.001$). We observed less differentiation (lower $F_{ST}$ and $R_{ST}$) among the three clusters predominated by the *P. vulgaris* accessions (C1, C2, C3), than between any of these and C4, which was comprised exclusively of *P. coccineus* accessions. When considering these comparisons with the *P. coccineus* cluster (C4), the lowest $F_{ST}$ was with...
FIGURE 2 | Genetic relationships within the whole set of accessions, as determined by principal component analysis. MW, Mesoamerican wild; AW, Andean wild; PhI, northern Peru and Ecuador.

FIGURE 3 | Genetic relationships within the *P. vulgaris* accessions, as assessed by principal component analysis. MW, Mesoamerican wild; AW, Andean wild; PhI, northern Peru and Ecuador; (A,B), major groups identified by PCA analysis.
Table 2 | Distribution of the accessions into the four cpSSR clusters (C1, C2, C3, C4) identified by the BAPS analysis.

| Accession | Cluster |
|-----------|---------|
|           | C1      | C2    | C3    | C4 |
| MW        | 13      | 21    | 21    | –  |
| AW        | 46      | –     | 1     | –  |
| Phi       | –       | 3     | 4     | –  |
| P. coccineus | –   | 2     | –     | 8  |
| Overall   | 59      | 26    | 26    | 8  |

MW, Mesoamerican wild; AW, Andean wild; Phi, northern Peru and Ecuador.

Table 3 | Genetic diversity estimates computed for the 17 cpSSRs considering the four clusters (C1, C2, C3, and C4) identified by BAPS analysis.

| Cluster | N | % polymorphic loci | Na | Ne | Np (freq. ≥ 0.05) | He |
|---------|---|---------------------|----|----|-------------------|----|
| C1      | 59| 88.2                | 3.4| 2.0| 6                 | 0.42|
| C2      | 26| 88.2                | 3.2| 2.1| 3                 | 0.45|
| C3      | 26| 88.2                | 3.1| 1.8| 7                 | 3  |
| C4      | 8 | 52.9                | 1.9| 1.6| 10                | 0.29|

N, sample size; Na, mean number of observed alleles per locus; Ne, mean effective number of alleles per locus; Np, number of private alleles; Np (freq. ≥ 0.05), number of private alleles with frequency higher than 0.05; He, expected heterozygosity.

Table 4 | Genetic divergence (FST and RST, below and above the diagonal, respectively) within the P. vulgaris populations and with P. coccineus.

|          | MW | AW | Phi | P. coccineus |
|----------|----|----|-----|--------------|
| MW       | –  | 0.13*| 0.08| 0.58**       |
| AW       | 0.24*| –   | 0.21**| 0.78**       |
| Phi      | 0.12| 0.70**| –   | 0.60**       |
| P. coccineus | 0.33**| 0.49**| 0.38**| –           |

Significance obtained by 10,000 permutations: **P ≤ 0.001; *P ≤ 0.01.

Table 5 | Genetic divergence (FST and RST, below and above the diagonal, respectively) between the four cpSSR clusters identified by population structure analysis.

|          | C1 | C2 | C3 | C4 |
|----------|----|----|----|----|
| C1       | –  | 0.54**| 0.37**| 0.90**|
| C2       | 0.26**| –   | 0.15**| 0.81**|
| C3       | 0.28**| 0.37**| –   | 0.68**|
| C4       | 0.50**| 0.39**| 0.56**| –   |

Significance obtained by 10,000 permutations: **P ≤ 0.001.

FIGURE 4 | Geographical distribution of the P. vulgaris accessions based on the BAPS cluster membership.

FIGURE 5A,B shows the geographical distribution of these clusters. The analysis of contingency tables indicated a significant association (P < 0.0001; likelihood ratio χ² test) between the genetic clusters obtained with these different markers (Figure 5C). In particular, cluster C1 was represented by clusters B3, B4, and B6, while cluster C2 included the B1, B2, and B5 clusters. In contrast, cluster C3 did not show any associations, although it is represented by accessions from the gene pools from Mesoamerica (B1, B2, B3), the Andes (B6), and northern Peru and Ecuador (B5).

DISCUSSION

The main aim of the present study was to investigate the complex evolutionary history that characterizes P. vulgaris through an analysis of its genetic diversity at the plastidial DNA level, in comparison with the study of Bitocchi et al. (2012) that was based

the C2 cluster [FST(C2-C4) = 0.39]. RST gave a slightly different pattern, with comparisons involving the C3 cluster showing the lowest RST (Table 5).

NUCLEOTIDE DATA VERSUS cpSSRs

The availability of sequence data for five gene fragments for 71 out of the 109 P. vulgaris accessions allowed a comparison between these different kinds of data (plastidial and nuclear). Three clusters were identified by the analysis carried out with cpSSRs. They corresponded to clusters (C1, C2, and C3) determined previously using all the accessions, while the Cluster C4 was not determined due to the exclusion, in this comparative analysis, of the P. coccineus accessions. Six clusters (B1, B2, B3, B4, B5, and B6), as in Bitocchi et al. (2012) were identified with nuclear nucleotide data. The

distribution of the accessions into the nucleotide data and cpSSR clusters is reported in Table 6. Figures 5A,B shows the geographical distribution of these clusters. The analysis of contingency tables indicated a significant association (P < 0.0001; likelihood ratio χ² test) between the genetic clusters obtained with these different markers (Figure 5C). In particular, cluster C1 was represented by clusters B3, B4, and B6, while cluster C2 included the B1, B2, and B5 clusters. In contrast, cluster C3 did not show any associations, although it is represented by accessions from the gene pools from Mesoamerica (B1, B2, B3), the Andes (B6), and northern Peru and Ecuador (B5).
Table 6 | Distribution of the 71 accessions shared between nucleotide and cpSSR data into the six nucleotide data clusters (B1, B2, B3, B4, B5, and B6) and the four cpSSR clusters (C1, C2, C3, C4) identified by the BAPS analysis.

| Accession | cpSSR cluster | Nucleotide data cluster |
|-----------|---------------|-------------------------|
|           | C1 | C2 | C3 | B1 | B2 | B3 | B4 | B5 | B6 |
| MW        | 7  | 15 | 4  | 12 | 7  | 5  | 2  | –  | –  |
| AW        | 40 | –  | 1  | –  | –  | –  | –  | –  | 41 |
| Phi       | –  | 3  | 1  | –  | –  | –  | –  | –  | 4  |
| Overall   | 47 | 18 | 6  | 12 | 7  | 5  | 2  | 4  | 41 |

MW, Mesoamerican wild; AW, Andean wild; Phi, northern Peru and Ecuador.

on nuclear nucleotide data. Thus, taking into account the specific characteristics of the plastidial genome (haploidy, lack of recombination, uniparental inheritance), we used cpSSRs to contribute to the existing knowledge of the evolution of the common bean and its closely related species, and to provide new insights, especially considering that comparisons of data obtained through analyses of nuclear and cytoplasmic DNA can provide contrasting information on evolutionary processes (Birky, 1988; McCauley, 1995; Ennos et al., 1999; Provan et al., 1999; Weising and Gardner, 1999; Ishii et al., 2001; Lira et al., 2003; Ueno et al., 2005).

The data obtained here are in agreement with the Mesoamerican origin of P. vulgaris, thus confirming the findings of Bitocchi et al. (2012), where the nucleotide diversity at five nuclear gene fragments in a wide sample of wild P. vulgaris accessions was analyzed (mostly shared with the present study). Moreover, the absence of phaselolin type I in the Mesoamerican gene pool might be due to its extinction in Mesoamerica, or it might still be present, but just not included in the samples analyzed in the literature.

The first outcome was the reduction in the genetic diversity (26%) in the Andean gene pool, compared to that of Mesoamerica. This has already been shown, even if to different extents, by analyses carried out with different nuclear molecular markers (SSRs: 7%, Kwak and Gepts, 2009; AFLPs: 45%, Rossi et al., 2009) and sequence data (90%, Bitocchi et al., 2012). In particular, the loss of diversity detected with cpSSRs is intermediate between the SSRs and AFLPs, as is their mutation rate ($10^{-5}$–$10^{-6}$ mutations per generation; Provan et al., 1999; Marshall et al., 2002). Indeed SSRS are characterized by a very high mutation rate ($10^{-3}$–$10^{-4}$ mutations per generation; Estoup and Angers, 1998; Mariette et al., 2001; Udupa and Baum, 2001; Viguouroux et al., 2002; Thuillet et al., 2005; Garoia et al., 2007) and AFLPs by a lower one ($10^{-6}$–$10^{-5}$ mutations per generation; Mariette et al., 2001; Gaudeul et al., 2004; Kropf et al., 2009). Consistent with the evidence obtained for the nuclear genome (Kwak and Gepts, 2009; Rossi et al., 2009; Nanni et al., 2011; Bitocchi et al., 2012), our data provide further evidence of the bottleneck that occurred before domestication of the common bean in the Andes, which led to impoverishment of the genetic diversity also at the plastidial level in the present gene pool. Moreover, this confirms the strong relationship between the mutation rate and the time needed for a population to recover the genetic diversity that can be lost after a bottleneck: the higher the mutation rate, the shorter the time needed (Glémin and Bataillon, 2009; Rossi et al., 2009; Nanni et al., 2011; Bitocchi et al., 2012, 2013).

Moreover, the BAPS analysis allows the division into three main clusters for the P. vulgaris accessions (C1, C2, C3). The Andean accessions are almost all included in cluster C1, with the only exception being an accession from southern Peru that belongs to cluster C3. Considering the nuclear data, cluster C1 is significantly associated with clusters B3, B6, and B4. This supports the close relationship between the Andean (B6) and the MW accessions from central Mexico (B3; Bitocchi et al., 2012), which indicates that these MW accessions represent the most probable plant material that spread and adapted to the southern part of the Andes.

Cluster C2 is characterized by the Mesoamerican accessions assigned using nucleotide data to clusters B1 and B2, and three of the seven Phi accessions, while cluster C3 groups the accessions that are representative of all of the gene pools (Mesoamerican, Andean, and northern Peru and Ecuador). These data provide further confirmation of the evidence highlighted by the nuclear data (Bitocchi et al., 2012); indeed, the Mesoamerican population is highly subdivided also at the plastidial level, and all of the genetic groups identified are present in particular in Central Mexico, which indicates this geographical area as the center of origin of P. vulgaris.

However, an interesting and novel outcome is revealed by the cpSSRs, which is probably due to the different characteristics of the nuclear and plastidial genome (and in particular to the presence of recombination for the nuclear genome): the identification of cluster C3 as a genetic group that incorporates accessions that are representative of all of the gene pools (MW, AW, Phi) and are not significantly associated with any genetic cluster identified with the nuclear data. In particular, almost all of the MW in cluster C3 are from Central Mexico, with the only exception being one Colombian genotype; moreover, cluster C3 comprised four Phi accessions and one AW accession. The wide distribution in cluster C3 can be interpreted as evidence that these accessions carry the ancestral plastidial genome that spread over the entire distribution that is now covered by P. vulgaris. This pattern is also confirmed by the $R_{ST}$ divergence estimations, where cluster C3 shows the lowest values compared to all of the other clusters, including most of the various alleles, when the size of the alleles is considered as a measure of the evolutionary distance among alleles. However, the same does not hold when the infinite allele model is considered: $F_{ST}$. Indeed, for $F_{ST}$, C2 shows the lowest divergence. This appears to be determined by the higher diversity (He) of C2 compared to C3, but not as alleles number (richness), with C2 showing the more uniform distribution of allele frequencies. Thus, we can speculate that the different results obtained for $R_{ST}$ and $F_{ST}$ might be the result of the more precise estimation of allele divergence using $R_{ST}$ and because C3 has more skewed allele frequencies due to the drift (e.g., a bottleneck).

The membership of the two P. coccineus genotypes to cluster C2 suggests that this cluster can be considered as having been derived from an ancestral lineage from which P. vulgaris separated from P. coccineus. Alternatively, this might result from post speciation introgression from P. vulgaris (with P. vulgaris as the maternal parent of the initial hybridization). This putative introgression of plastidial DNA from P. vulgaris to P. coccineus is consistent
with the hypothesis that the *P. dumosus* species originated from a cross of *P. vulgaris* as maternal and *P. coccineus* as paternal parent, followed by successive backcrosses from *P. coccineus* as paternal donor (Schmit et al., 1993; Ila et al., 1994; Angioi et al., 2009a). Indeed, *P. dumosus* is closer to *P. coccineus* according to nuclear DNA comparisons (Piñero and Eguiarte, 1988; Delgado-Salinas...
et al., 1999), while according to chloroplast DNA comparisons it appears to be more closely related to P. vulgaris (Llaca et al., 1994; Angioi et al., 2009a). These outcomes reveal the complexity of the evolution of P. vulgaris within the evolutionary history of its closely related species, P. coccineus and P. dumosus (Schmit et al., 1993; Delgado-Salinas et al., 1999, 2006; Chacon et al., 2007), both of which are found in Mesoamerica (Schmit and Debouck, 1991; Freytag and Debouck, 2002). In spite of the marked differences in mating systems and life cycles, P. coccineus (predominantly alloagamous and perennial), P. vulgaris (predominantly autogamous and annual), and P. dumosus (intermediate characteristics between P. coccineus and P. vulgaris) are partially intercrossable, although only when P. vulgaris is the female parent (Mendel, 1866; Wall, 1970; Shi et al., 1982; Huc and Scoles, 1985). However, further studies should be carried out here, to compare a larger sample that includes genotypes from all three of these sister species and uses both nuclear and plastidial DNA analyses.

CONCLUSION

Chloroplast SSRs are widely used for evolutionary and phylogenetic studies as they have been demonstrated to be effective indicators of the genetic structure of a population. Therefore, we used this alternative form of analysis (with respect to nuclear data) with the aim of obtaining a more detailed picture of the history of the common bean. These cpSSR data strongly support the nuclear data of Bitocchi et al. (2012), that indicated a clear Mesoamerican origin of this species, and in particular, they support Central Mexico as, with high probability, the cradle of common bean diversity.

A novel outcome was also provided by these analyses based on the polymorphism at the chloroplast DNA level: the identification of a genetic group (cluster C3) that includes accessions distributed from northern Mexico to Peru that appear to carry a putative ancestral plastidial genome.

Finally, the present study highlights the potential to evaluate the evolutionary history of P. vulgaris within the evolution of the whole species complex that includes P. vulgaris, P. coccineus, and P. dumosus. A deeper study of the formation and evolution of these closely related and intercrossable species will be intriguing from an evolutionary point of view. At the same time, such data should be particularly relevant for common bean breeding programs, as demonstrated by the increasing interest in the development of interspecific lines (P. vulgaris-P. coccineus and P. vulgaris-P. dumosus crosses) for the introgression of important traits; e.g., resistance to biotic and abiotic stress in P. vulgaris elite germplasm (Singh et al., 2009; Klaedtke et al., 2012).

ACKNOWLEDGMENTS

This work was supported by grants from the Italian Government (MIUR; Grant number 20083PFSX_001, PRIN Project 2008), the Università Politecnica delle Marche (2008–2011) and the Marche Region (Grant number L.R.37/99 art. 2let. I – PARDGR 24710 – DDPFP98/CS10).

REFERENCES

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.

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.

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.

Becerra-Velásquez, V. L., and Gepts, P. (1994). RFLP diversity in common bean (Phaseolus vulgaris L.). Genome 37, 256–263.

Birky, C. W. (1988). “Evolution and population genetics of organelle genome mechanisms and models,” in Plant Evolutionary Biology, eds L. D. Gottlieb, and S. K. Jain (London: Chapman and Hall), 23–53.

Bitocchi, E., Nanni, L., Bellucci, E., Rossi, M., Giardini, A., Spagnoletti Zeuli, P., et al. (2012). Mesoamerican origin of the common bean (Phaseolus vulgaris L.) is revealed by sequence data. Proc. Natl. Acad. Sci. USA. 109, E788–E796.

Bitocchi, E., Bellucci, E., Giardini, G., Rau, R., 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, 500–513.

Chacon, S. M. I., Pickersgill, B., Debouck, D. G., and Salvador Arias, J. (2007). Phylegographic analysis of the chloroplast DNA variation in wild common bean (Phaseolus vulgaris L.) in the Americas. Plant Syst. Evol. 266, 175–195.

Chung, S., and Staub, J. E. (2003). The development and evaluation of consensus chloroplast primer pairs that possess highly variable sequence regions in a diverse array of plant taxa. Theor. Appl. Genet. 107, 757–767.

Corander, J., Marttinen, P., Sirén, J., and Tang, J. (2008). Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinformatics 9:539. doi:10.1186/1471-2105-9-539

Corander, J., Waldmann, P., and Siljanpää, M. J. (2003). Bayesian analysis of genetic differentiation between populations. Genetics 163, 367–374.

Debouck, D. G., Toro, O., Paredes, O. M., Johnson, W. C., and Gepts, P. (1993). “The wild relative of Phaseolus vulgaris in Mesoamerica and the Andes.” in Theor. Appl. Genet. 107, 757–767.

Ennos, R. A., Sinclair, W. T., Hu, X. S., and Langdon, A. (1999). “Using organelle markers to elucidate the history, ecology and evolution of plant populations.” in Molecular Systematics and Plant Evolution, eds P. M. Hollingsworth, R. B. Bateman, and R. J. Gorrell (London: Taylor & Francis), 1–19.

Estoup, A., and Angers, B. (1998). “Microsatellites and minisatellites for molecular ecology: theoretical and experimental considerations,” in Advances in Molecular Ecology, ed G. Carvalho (Amsterdam: IOS Press), 55–86.

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite version 3.5: A new series of programs for population genetics analyses under Linux and windows. Mol. Ecol. Resour. 10, 564–567.

Freytag, G. E., 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.

Gariofa, C., Guarniero, I., Grifoni, D., Marzola, S., and Tinti, F. (2007). Comparative analysis of AFLPs and SSRs efficiency in resolving population genetic structure of Mediterranean Sisola vulgaris. Mol. Ecol. 16, 1377–1387.

Gaudet, M., Till-Bottraud, I., Barjon, F., and Manel, S. (2004). Genetic diversity and differentiation in Eryngium alpinum L. (Apiaceae): Comparison of AFLP and microsatellite markers. Heredity 92, 508–518.

Gepts, P., and Debouck, D. G. (1991). “Origin, domestication, and evolution of the common bean, Phaseolus vulgaris,” in Common Beans: Research for Crop Improvement, eds O. Voyset and A. Van Schoonhoven (Wallingford, Oxon: CAB International), 7–53.
Désiderio et al.

Frontiers in Plant Science | Plant Genetics and Genomics
January 2013 | Volume 3 | Article 312 | 10

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

Gleim, S., and Bataillon, T. (2009). A comparative view of the evolution of grasses under domestication. New Phytol. 183, 273–290.

Holsinger, K. E., and Weir, B. S. (2009). Genetics in geographically structured populations: defining, estimating and interpreting FST. Nat. Rev. Genet. 10, 639–650.

Hougaard, B. K., Heegaard Madsen, L., Sandal, N., de Carvalho Moretzsohn, M., Fredlund, J., Schauer, L., et al. (2008). Legume anchor markers link syntenic regions between Phaseolus vulgaris, Lotus japonicus, Medicago truncatula and Arachis. Genetics 179, 2299–2312.

Hucl, P., and Scoles, G. J. (1985). Inter- species hybridization in the common bean: a review. Hort. Sci. 20, 352–357.

Ishii, T., Xu, Y., and McCouch, S. R. (2012). Photosynthate remobilization of presumed ancestral DNA in commercial or financial relationships that could be perceived as a conflict of interest.

Received: 30 October 2012; paper pending published: 27 November 2012; accepted: 27 December 2012; published online: 22 January 2013.

Citation: Désiderio D, Bitocchi E, Bellucci E, Rau D, Rodriguez M, Attenge T, Papa R and Nanni L (2013) Chloroplast microsatellite diversity in Phaseolus vulgaris. Frontiers in Plant Science, 4:312. doi: 10.3389/fpls.2012.00312 This article was submitted to Frontiers in Plant Genetics and Genomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Desiderio, Bitocchi, Bellucci, Rau, Rodriguez, Attenge, Papa R and Nanni L. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.
APPENDIX

Table A1 | List of accessions used in this study.

| Accession code | Synonyms | Species | Donor | Pop code | Country | Department, state, or province | Latitude | Longitude | BAPS cluster (cpSSR) | BAPS cluster (nucleotide data); $q \geq 0.6$ |
|----------------|----------|---------|-------|----------|---------|-------------------------------|----------|-----------|---------------------|---------------------------------|
| G21113         | LEROI COL-14, NI-922 | *P. vulgaris* | CIAT | MW | Colombia | Cundinamarca | 44,833 | −73,933 | C2 | / |
| G22304         | LEROI COL-13, NI-1142 | *P. vulgaris* | CIAT | MW | Colombia | Cundinamarca | 44,833 | −73,933 | C3 | / |
| G21115         | LEROI COL-23, NI-926 | *P. vulgaris* | CIAT | MW | Colombia | Cundinamarca | 45,333 | −73,917 | C2 | B1 |
| G21117         | LEROI COL-28, NI-837 | *P. vulgaris* | CIAT | MW | Colombia | Cundinamarca | 46,667 | −74.4 | C2 | B1 |
| G22303         | LEROI COL-22, NI-1141 | *P. vulgaris* | CIAT | MW | Colombia | Cundinamarca | 45,333 | −73,917 | C1 | B1 |
| G23462         | LEROI COL-15, NI-1256, X-636 | *P. vulgaris* | CIAT | MW | Colombia | Cundinamarca | 50,833 | −73,817 | C2 | B1 |
| G2771          | GENTRY 22274; PI318702 | *P. vulgaris* | CIAT | MW | Mexico | Nayarit | 211,667 | −104.37 | C3 | / |
| G11051         | DGD-451 | *P. vulgaris* | CIAT | MW | Mexico | Jalisco | 207,667 | −103.4 | C3 | / |
| G12927         | M7278-G, Pi417689 | *P. vulgaris* | CIAT | MW | Mexico | Jalisco | 20.7 | −102.35 | C3 | / |
| G12957         | M7424-C, Pi417786 | *P. vulgaris* | CIAT | MW | Mexico | Jalisco | 20.9 | −102.37 | C1 | / |
| G23418         | DGD-2111 | *P. vulgaris* | CIAT | MW | Costa Rica | San Jose | 98,667 | −84.117 | C2 | / |
| G23558         | OAXACA 112 | *P. vulgaris* | CIAT | MW | Mexico | Oaxaca | 163,333 | −95.233 | C2 | / |
| G24366         | JSG & LOS-150 | *P. vulgaris* | CIAT | MW | Mexico | Jalisco | 204,833 | −103.4 | C1 | / |
| Pi417775       | G12949, M7408-P | *P. vulgaris* | CIAT | MW | Mexico | Jalisco | 20.64 | −102.41 | C3 | / |
| W812107        | CR-93-004 | *P. vulgaris* | USDA | MW | Costa Rica | Puntarenas | 8.95 | −83.038 | C2 | / |
| G9989          | HM7395-BULK | *P. vulgaris* | CIAT | MW | Mexico | Jalisco | 20.5 | −104.82 | C3 | B1 |
| G19906         | DGD-1610 | *P. vulgaris* | CIAT | MW | Guatemala | Sacatepequez | 14.45 | −90.7 | C2 | B1 |
| G19907         | DGD-1611 | *P. vulgaris* | CIAT | MW | Guatemala | Sacatepequez | 14.45 | −90.817 | C2 | B1 |
| G19909         | DGD-1619 | *P. vulgaris* | CIAT | MW | Guatemala | Sacatepequez | 14.55 | −90.833 | C2 | B1 |
| G22837         | GN 84127/BB 8480, P16-001 | *P. vulgaris* | CIAT | MW | Mexico | Chihuahua | 269,333 | −106.42 | C2 | B1 |
| G23463         | GN 84154, L 625 | *P. vulgaris* | CIAT | MW | Mexico | Chihuahua | 283,333 | −108.5 | C2 | B1 |
| G24378         | JSG & LOS-199 | *P. vulgaris* | CIAT | MW | Mexico | Oaxaca | 16.4 | −97.083 | C2 | B1 |
| G50899         | LEROI MEX-26, NI-1144 | *P. vulgaris* | CIAT | MW | Mexico | Durango | 237,833 | −105.37 | C2 | B1 |
| G11056         | DGD-490 | *P. vulgaris* | CIAT | MW | Mexico | Jalisco | 205,667 | −104.77 | C3 | B2 |
| G20515         | M81378-1 | *P. vulgaris* | CIAT | MW | Mexico | Puebla | 19.8 | −97.783 | C2 | B2 |
| G23429         | DGD-2325 | *P. vulgaris* | CIAT | MW | Mexico | Puebla | 189,667 | −98.383 | C2 | B2 |
| G24571         | JSMM-4002 | *P. vulgaris* | CIAT | MW | Mexico | Oaxaca | 171,667 | −97.963 | C2 | B2 |
| G24572         | JSMM-4006 | *P. vulgaris* | CIAT | MW | Mexico | Oaxaca | 159,833 | −96.517 | C3 | B2 |
| G24599         | JAG-180 | *P. vulgaris* | CIAT | MW | Mexico | Chiapas | 164,833 | −92.517 | C2 | B2 |
| G50415         | JAG-209 | *P. vulgaris* | CIAT | MW | Mexico | Hidalgo | 20.85 | −98.717 | C2 | B2 |
| G11050         | DGD-439 | *P. vulgaris* | CIAT | MW | Mexico | Michoacan | 196,833 | −101.27 | C1 | B4 |
| G12922         | M7278-A, Pi417683 | *P. vulgaris* | CIAT | MW | Mexico | Jalisco | 20.7 | −102.35 | C3 | B3 |
| G12979         | M7439T | *P. vulgaris* | CIAT | MW | Mexico | Jalisco | 201,167 | −104.37 | C1 | B3 |
| G23415A        | DGD-2077 | *P. vulgaris* | CIAT | MW | Mexico | Queretaro | 211,333 | −99.617 | C1 | B3 |
| G23652         | M2058 | *P. vulgaris* | CIAT | MW | Mexico | Puebla | 19.8 | −97.783 | C1 | B3 |

(Continued)
Table A1 | Continued

| Accession code | Synonyms | Species | Donor | Pop code | Country | Department, Latitude state, or province | Longitude | BAPS cluster (cpSSR) | BAPS cluster (nucleotide data); q ≥ 0.6
|----------------|---------|---------|-------|----------|---------|----------------------------------------|----------|---------------------|---------------------------------|
| G12865 ●       | GENTRY 22199, PI3186996 | *P. vulgaris* | CIAT | MW | Mexico | Jalisco | 193,333 | −103.25 | C1 | B3 |
| G12873 ●       | PI325678, GENTRY22492 | *P. vulgaris* | CIAT | MW | Mexico | Morelos | 19 | −99.25 | C1 | B4 |
| G10012         | MORELOS 646, V-1434 | *P. vulgaris* | CIAT | MW | Mexico | Morelos | 188,833 | −99.15 | C2 | / |
| G12872         | GENTRY 22404, PI325677 | *P. vulgaris* | CIAT | MW | Mexico | Morelos | 189,667 | −99.1 | C1 | / |
| G12877         | GENTRY 22930, PI325683 | *P. vulgaris* | CIAT | MW | Mexico | Morelos | 18.95 | −99.217 | C3 | / |
| G12896         | M7230, PI417629 | *P. vulgaris* | CIAT | MW | Mexico | Michoacan | 201,333 | −102.08 | C3 | / |
| G12924         | M7278-C, PI417685 | *P. vulgaris* | CIAT | MW | Mexico | Jalisco | 20.7 | −102.35 | C3 | / |
| G12930         | M7278-L, PI417692 | *P. vulgaris* | CIAT | MW | Mexico | Jalisco | 20.7 | −102.35 | C3 | / |
| G13018         | MORELOS 654, V-1438 | *P. vulgaris* | CIAT | MW | Mexico | Morelos | 188,833 | −99.15 | C1 | / |
| G13055         | MORELOS 635, NI-404 | *P. vulgaris* | CIAT | MW | Mexico | Morelos | 188,833 | −99.15 | C2 | / |
| G12866         | GENTRY 22202, PI318697 | *P. vulgaris* | USDA | MW | Mexico | Jalisco | 19,683 | −103.48 | C1 | / |
| CHWENN2        |                     | *P. vulgaris* | UNIVPM | MW | Mexico | Chiapas | 164,833 | −92.517 | C3 | / |
| CHWITE16       |                     | *P. vulgaris* | UNIVPM | MW | Mexico | Chiapas | 164,833 | −92.517 | C3 | / |
| DGW15          |                     | *P. vulgaris* | UNIVPM | MW | Mexico | Durango | 237,833 | −105.37 | C3 | / |
| 111d           |                     | *P. vulgaris* | UNIVPM | MW | Mexico | Chiapas | 164,833 | −92.517 | C3 | / |
| JAL97          |                     | *P. vulgaris* | UNIVPM | MW | Mexico | Jalisco | 204,833 | −103.4 | C3 | / |
| MOW5           |                     | *P. vulgaris* | UNIVPM | MW | Mexico | Morelos | 189,667 | −99.1 | C1 | / |
| MXW17          |                     | *P. vulgaris* | UNIVPM | MW | Mexico | – | *** | *** | C3 | / |
| PUA21          |                     | *P. vulgaris* | UNIVPM | MW | Mexico | Puebla | – | *** | *** | C3 | / |
| G23415 ●       | DGD-2077 | *P. vulgaris* | CIAT | MW | Mexico | Querétaro | 211,333 | −99.617 | C1 | / |
| G23423C        | DGD-2157 | *P. vulgaris* | CIAT | AW | Perú | Apurímac | −13.85 | −72.967 | C1 | / |
| W617481        | P1638874 | *P. vulgaris* | USDA | AW | Argentina | Jujuy | −22.267 | −64.683 | C1 | / |
| W617500        | P1640966 | *P. vulgaris* | USDA | AW | Argentina | Salta | −24.65 | −65.367 | C1 | / |
| W617501        | P1640967 | *P. vulgaris* | USDA | AW | Argentina | Salta | −24.65 | −65.367 | C1 | / |
| G7225          | APURIMAC 76 | *P. vulgaris* | CIAT | AW | Perú | Apurímac | −13.667 | −72.883 | C1 | / |
| W617467        | P1638865 | *P. vulgaris* | USDA | AW | Argentina | Tucumán | −26.217 | −65.527 | C1 | / |
| G4699 ●        | NI-029 | *P. vulgaris* | CIAT | AW | Argentina | – | *** | *** | C3 | / |
| G10024 ●       | NI-190 | *P. vulgaris* | CIAT | AW | Argentina | – | *** | *** | C1 | B6 |
| G12856 ●       | P1260405, SMITH PV-1 | *P. vulgaris* | CIAT | AW | Perú | Huanuco | −10,333 | −76.183 | C3 | B6 |
| G19888 ●       | DGD-623 | *P. vulgaris* | CIAT | AW | Argentina | Jujuy | −24.167 | −65.6 | C1 | B6 |
| G19889 ●       | DGD-624 | *P. vulgaris* | CIAT | AW | Argentina | Jujuy | −24.25 | −65.283 | C1 | B6 |
| G19891 ●       | DGD-628 | *P. vulgaris* | CIAT | AW | Argentina | Salta | −25.117 | −65.617 | C1 | B6 |
| G19892 ●       | DGD-629 | *P. vulgaris* | CIAT | AW | Argentina | Salta | −25.15 | −65.65 | C1 | B6 |
| G19893 ●       | DGD-630, NEEMA S-211/S-226 | *P. vulgaris* | CIAT | AW | Argentina | Salta | −24,633 | −65,483 | C1 | B6 |
| G19895 ●       | DGD-637, NEEMA T-711/T-717 | *P. vulgaris* | CIAT | AW | Argentina | Tucumán | −26,433 | −65,517 | C1 | B6 |

(Continued)
Table A1 | Continued

| Accession code¹ | Synonyms | Species | Donor² | Pop code³ | Country | Department, state, or province | Latitude | Longitude | BAPS cluster (cpSSR) | BAPS cluster (nucleotide data); q ≥ 0.6⁴ |
|-----------------|----------|---------|--------|-----------|---------|--------------------------------|----------|-----------|---------------------|--------------------------------------------|
| G19896          | DGD-639  | P. vulgaris | CIAT   | AW       | Argentina | Tucuman | −26,217 | −65,583 | C1                  | B6                                         |
| G19897          | DGD-643, NEEMA T911/T917 | P. vulgaris | CIAT   | AW       | Argentina | Tucuman | −27,317 | −65,917 | C1                  | B6                                         |
| G19898          | DGD-644  | P. vulgaris | CIAT   | AW       | Argentina | Tucuman | −27,333 | −65,95  | C1                  | B6                                         |
| G19901          | DGD-649  | P. vulgaris | CIAT   | AW       | Argentina | Tucuman | −26,933 | −65,7   | C1                  | B6                                         |
| G21194          | DGD-621  | P. vulgaris | CIAT   | AW       | Argentina | Jujuy    | −24,117 | −65,417 | C1                  | B6                                         |
| G21197          | DGD-1711 | P. vulgaris | CIAT   | AW       | Argentina | Jujuy    | −24,05  | −65,45  | C1                  | B6                                         |
| G21198          | DGD-1712 | P. vulgaris | CIAT   | AW       | Argentina | Jujuy    | −24,067 | −65,367 | C1                  | B6                                         |
| G21199          | DGD-1713 | P. vulgaris | CIAT   | AW       | Argentina | Jujuy    | −23,917 | −65,35  | C1                  | B6                                         |
| G21201          | DGD-1716 | P. vulgaris | CIAT   | AW       | Argentina | Salta    | −22,25  | −65     | C1                  | B6                                         |
| G23420          | DGD-2147 | P. vulgaris | CIAT   | AW       | Perú      | Junin    | −11,2   | −75,483 | C1                  | B6                                         |
| G23421          | DGD-2152 | P. vulgaris | CIAT   | AW       | Perú      | Junin    | −12,017 | −74,883 | C1                  | B6                                         |
| G23422          | DGD-2156 | P. vulgaris | CIAT   | AW       | Perú      | Apurimac | −14     | −73,167 | C1                  | B6                                         |
| G23426          | DGD-2295 | P. vulgaris | CIAT   | AW       | Perú      | Apurimac | −13,617 | −73,2   | C1                  | B6                                         |
| G23444          | DGD-2497 | P. vulgaris | CIAT   | AW       | Bolivia   | Chuquisaca | −19,3   | −64,317 | C1                  | B6                                         |
| G23445          | DGD-2501 | P. vulgaris | CIAT   | AW       | Bolivia   | Tarija    | −21,533 | −64,867 | C1                  | B6                                         |
| G23456          | DGD-2581 | P. vulgaris | CIAT   | AW       | Perú      | Cuzco    | −13,5   | −72,483 | C1                  | B6                                         |
| W61746          | PI638864 | P. vulgaris | USDA   | AW       | Argentina | Tucuman  | −26,233 | −65,483 | C1                  | B6                                         |
| W617468         | PI638866 | P. vulgaris | USDA   | AW       | Argentina | Tucuman  | −27,817 | −65,783 | C1                  | B6                                         |
| W617469         | PI638867 | P. vulgaris | USDA   | AW       | Argentina | Tucuman  | −27,797 | −65,785 | C1                  | B6                                         |
| W617470         | PI640964 | P. vulgaris | USDA   | AW       | Argentina | Tucuman  | −26,383 | −65,467 | C1                  | B6                                         |
| W617471         | PI638868 | P. vulgaris | USDA   | AW       | Argentina | Tucuman  | −26,383 | −65,533 | C1                  | B6                                         |
| W617472         | PI638869 | P. vulgaris | USDA   | AW       | Argentina | Tucuman  | −26,95  | −65     | C1                  | B6                                         |
| W617473         | PI638870 | P. vulgaris | USDA   | AW       | Argentina | Salta    | −26,1   | −65,6   | C1                  | B6                                         |
| W617474         | PI640965 | P. vulgaris | USDA   | AW       | Argentina | Salta    | −25,161 | −65,611 | C1                  | B6                                         |
| W617475         | PI638871 | P. vulgaris | USDA   | AW       | Argentina | Salta    | −25,167 | −65,617 | C1                  | B6                                         |
| W617476         | PI638872 | P. vulgaris | USDA   | AW       | Argentina | Salta    | −25,166 | −65,649 | C1                  | B6                                         |
| W617478         | PI638873 | P. vulgaris | USDA   | AW       | Argentina | Salta    | −24,896 | −65,801 | C1                  | B6                                         |
| W617486         | PI638875 | P. vulgaris | USDA   | AW       | Argentina | Jujuy    | −22,267 | −64,683 | C1                  | B6                                         |
| W617489         | PI661807 | P. vulgaris | USDA   | AW       | Argentina | Salta    | −24,9   | −65,483 | C1                  | B6                                         |
| W617502         | PI640968 | P. vulgaris | USDA   | AW       | Argentina | Salta    | −24,717 | −65,483 | C1                  | B6                                         |
| W618821         | PI638897 | P. vulgaris | USDA   | AW       | Bolivia   | Chuquisaca | −19,283 | −64,333 | C1                  | B6                                         |
| W618826         | PI638898 | P. vulgaris | USDA   | AW       | Bolivia   | Chuquisaca | −19,283 | −64,333 | C1                  | B6                                         |
| G23581          | DGD-2765 | P. vulgaris | CIAT   | Phi      | Ecuador   | Azuay    | −3,2    | −79,183 | C3                  | /                                          |
| G23582          | DGD-2769 | P. vulgaris | CIAT   | Phi      | Ecuador   | Chimborazo | −22,667 | −78,967 | C3                  | /                                          |
| G23724          | DGD-2881 | P. vulgaris | CIAT   | Phi      | Ecuador   | Loja   | −43,167 | −79,933 | C3                  | /                                          |
| G21246          | DGD-1962 | P. vulgaris | CIAT   | Phi      | Perú      | Cajamarca | −71,167 | −78,783 | C2                  | B5                                         |
| G23585          | DGD-2855 | P. vulgaris | CIAT   | Phi      | Perú      | Cajamarca | −6,35   | −79,4   | C2                  | B5                                         |
| G23587          | DGD-2858 | P. vulgaris | CIAT   | Phi      | Perú      | Cajamarca | −6,35   | −79,4   | C3                  | B5                                         |
| G23726          | DGD-2889 | P. vulgaris | CIAT   | Phi      | Ecuador   | Chimborazo | −19,667 | −78,95  | C2                  | B5                                         |
| PI553280        | TARS212, 78-G-4 | P. coccineus | USDA   | –        | Guatemala | Sacatepequez | 14,43   | −90,95  | C2                  | /                                          |
| PI553287        | TARS222, 78-G-15 | P. coccineus | USDA   | –        | Guatemala | Sacatepequez | 14,67   | −90,75  | C2                  | /                                          |
| PI325584        | ACAHUATE | P. coccineus | USDA   | –        | Mexico    | Puebla   | 19,816  | −978,166 | C4                  | /                                          |
| PI417608        | M7417-G | P. coccineus | USDA   | –        | Mexico    | Jalisco   | 20,866  | −102,367 | C4                  | /                                          |

(Continued)
Table A1 | Continued

| Accession code | Synonyms | Species | Donor | Pop code | Country | Department, Latitude state, or province | Longitude | BAPS cluster (cpSSR) | BAPS cluster (nucleotide data); q ≥ 0.6 |
|----------------|----------|---------|-------|----------|---------|-----------------------------------------|----------|----------------------|------------------------------------------|
| PI417611       |          | P. coccineus | USDA – Mexico Jalisco | 20,866 | −102,366 | C4 | / |
| PI417592       |          | P. coccineus | USDA – Mexico Jalisco | 25.56 | −106.37 | C4 | / |
| PI430191       |          | P. coccineus | USDA – Mexico Chihuahua | 28.6 | −107,167 | C4 | / |
| PI430192       |          | P. coccineus | USDA – Mexico Chihuahua | 28.6 | −107,167 | C4 | / |
| CX 03          |          | P. coccineus | USDA – Mexico Morelos | *** ** | C4 | / |
| CF19           |          | P. coccineus | USDA – Mexico Morelos | *** ** | C4 | / |

1 Population code: WM, wild Mesoamerican; WA, wild Andean; PhI, Phaseolin I type.
2 •(dot) indicates the P. vulgaris accessions shared with the study of Bitocchi et al. (2012), showing high-quality sequences for all of the five Leg markers analyzed in Bitocchi et al. (2012); these accessions were used to compare the population structure results obtained using both cpSSR and nucleotide data.
3 CIAT, International Centre for Tropical Agriculture; USDA, United States Department of Agriculture; UNIVPM, Università Politecnica delle Marche.
4 q, percentage of membership to one of the six clusters identified using nucleotide data; a q threshold value of 0.6 was considered to assign accessions to clusters.

Table A2 | List of SSR used in this study.

| Locus | Primer sequence 5′−3′ | PCR conditionsa | Reference |
|-------|-----------------------|-----------------|-----------|
| ccSSR2 | fw-AATCCTGGACGTGAGAAATAA rev-AATCCCTCTCTCTTCCGTGTA | 1 Chung and Staub (2003) |
| ccSSR4 | fw-AGGTCAAAAATCCTATGAGGCA rev-TTTTTTGAAGAAGCTATTCARGAAC | 1 Chung and Staub (2003) |
| ccSSR7 | fw-CCCGAAGGCTGGCTCGGCA rev-GTCCGAGCTTCTCTCTCTTTT | 1 Chung and Staub (2003) |
| ccSSR8 | fw-TGATCTTACCTGGTCTCTCTCTA rev-TACATTACGTGAGATCTCC | 1 Chung and Staub (2003) |
| ccSSR9 | fw-GAGGATACACGACAGARGGARTTG rev-CCATACAGATGAGTGATGTTA | 1 Chung and Staub (2003) |
| ccSSR11 | fw-TGCTGACTCTAACCCTCCC | 2 Chung and Staub (2003) |
| ccSSR12 | fw-CAAAAAACCTTGAGATCCTAACACTC rev-TCTCAATAGCTGATCAGTTTA | 1 Chung and Staub (2003) |
| ccSSR15 | fw-GTCTGATTCCCTCCCTCTATG rev-TGCAATAGCTGATCATTATA | 1 Chung and Staub (2003) |
| ccSSR16 | fw-TACGGATGATCCCTCTCTTCTT rev-TCTGCAATAGCTGATCAGTTA | 1 Chung and Staub (2003) |
| ccSSR18 | fw-TCTGGAAGATCTTCTCTCTTCTCT rev-TCTGCAATAGCTGATCAGTTA | 1 Chung and Staub (2003) |
| ccSSR19 | fw-TACTGCAATAGCTTCTTCTCTCTCTCT rev-TCTGCAATAGCTGATCAGTTA | 1 Chung and Staub (2003) |
| ccSSR20 | fw-CCGCARATATGGAAAAACWAACAA rev-GCTAARCAATATGCTTCTCC | 1 Chung and Staub (2003) |
| cccmp2 | fw-GATCCCGGAGTGTAATCTCT rev-ATGGTAGAGAGGCTTCCAAGT | 1 Weising and Gardner (1999) |
| cccmp3 | fw-CAGACAAAAAGCTGACATAG rev-GTTCTATGCGCCGCTTCTTTAT | 3 Weising and Gardner (1999) |
| cp1 | fw-GATAATCAAAGAGCGATTAGG rev-GTCACCCCAGGAGCGTAA | 1 Angioi et al. (2009a) |
| cp2 | fw-TCTGTTTTTGACATATAGC rev-GTCATAAAAAAGGACGAAAAA | 4 Angioi et al. (2009a) |
| cp3 | fw-TGGGTGATCAATGATAAAAACGAAA rev-GCTGGCTAGGAAAGACTCTATAGGAA | 4 Angioi et al. (2009a) |

1 PCR conditions: 1: 5 min at 94˚C; 35 cycles of 1′ at 94˚C, 1′ at 50˚C, 1′ at 72˚C; 35 at 94˚C; 2: 5′ at 94˚C; touch down cycles 53–45˚C with −1˚C/2 cycles, 1′ at 72˚C, 20 cycles of 1′ at 94˚C, 1′ at 45˚C 1′ at 72˚C; 35 at 94˚C; 3: 5′ at 94˚C; touch down cycles 53–43˚C with −1˚C/2 cycles, 1′ at 72˚C; 20 cycles of 1′ at 94˚C 1′ at 43˚C, 1′ at 72˚C, 35 at 94˚C, 4: 5′ at 94˚C, 30 cycles of 30 s at 94˚C 30 s at 46˚C 30 s at 72˚C; 35 at 72˚C.
Table A3 | Number of alleles (Na) and gene diversity (He, Nei, 1978) in the overall, *P. vulgaris* and *P. coccineus* samples for each of the 17 cpSSRs used.

| Locus     | cp1 | cp2 | cp3 | ccmp2 | ccmp3 | ccSSR2 | ccSSR4 | ccSSR7 | ccSSR8 | ccSSR9 | ccSSR11 | ccSSR12 | ccSSR15 | ccSSR16 | ccSSR18 | ccSSR19 | ccSSR20 |
|-----------|-----|-----|-----|-------|-------|--------|--------|--------|--------|--------|---------|---------|---------|---------|---------|---------|---------|
| OVERALL   |     |     |     |       |       |        |        |        |        |        |         |         |         |         |         |         |         |
| He        | 0.48| 0.13| 0.15| 0.68  | 0.52  | 0.66   | 0.71   | 0.65   | 0.61   | 0.79   | 0.51    | 0.60    | 0.33    | 0.46    | 0.44    | 0.85    |
| Na        | 4   | 2   | 6   | 4     | 4     | 6      | 7      | 6      | 4      | 9      | 3       | 3       | 6       | 3       | 12      |         |
| Allelic   | 111–114 | 180–183 | 154–171 | 196–199 | 79–94 | 167–170 | 244–249 | 299–308 | 224–265 | 133–136 | 164–183 | 203–206 | 262–264 | 353–356 | 260–266 | 376–378 | 312–324 |
| *P. vulgaris* |     |     |     |       |       |        |        |        |        |        |         |         |         |         |         |         |         |
| He        | 0.43| 0.00| 0.02| 0.67  | 0.48  | 0.65   | 0.66   | 0.64   | 0.60   | 0.76   | 0.44    | 0.61    | 0.30    | 0.49    | 0.46    | 0.84    |
| Na        | 4   | 1   | 2   | 4     | 3     | 4      | 5      | 6      | 3      | 8      | 2       | 3       | 3       | 6       | 3       | 11      |         |
| Allelic   | 111–114 | 180 | 170–171 | 196–199 | 83–94 | 167–170 | 245–249 | 299–308 | 224–265 | 133–135 | 164–174 | 205–206 | 262–264 | 353–356 | 260–266 | 376–378 | 314–324 |
| *P. coccineus* |     |     |     |       |       |        |        |        |        |        |         |         |         |         |         |         |         |
| He        | 0.00| 0.36| 0.82| 0.00  | 0.64  | 0.36   | 0.36   | 0.64   | 0.00   | 0.64   | 0.64    | 0.36    | 0.00    | 0.53    | 0.00    | 0.00    | 0.84    |
| Na        | 1   | 2   | 5   | 1     | 3     | 2      | 2      | 4      | 1      | 4      | 3       | 2       | 1       | 2       | 1       | 1       | 6       |
| Alleles (bp) | 112 | 180–183 | 154–170 | 196 | 79–84 | 167–168 | 244–245 | 303–307 | 260 | 133–136 | 164–183 | 203–206 | 263 | 353–354 | 263 | 376 | 312–320 |