

RESEARCH ARTICLE

Morphological and molecular characterization of variation in common bean (*Phaseolus vulgaris* L.) germplasm from Azad Jammu and Kashmir, Pakistan

Sidra Nasar¹, Kate Ostevik^{2,3}, Ghulam Murtaza¹, Mark D. Rausher^{2*}

1 Department of Botany, University of Azad Jammu and Kashmir, Muzaffarabad, Pakistan, **2** Department of Biology, Duke University, Durham, North Carolina, United States of America, **3** Department of Evolution, Ecology, and Organismal Biology, University of California Riverside, Riverside, California, United States of America

* mrausher@duke.edu



OPEN ACCESS

Citation: Nasar S, Ostevik K, Murtaza G, Rausher MD (2022) Morphological and molecular characterization of variation in common bean (*Phaseolus vulgaris* L.) germplasm from Azad Jammu and Kashmir, Pakistan. PLoS ONE 17(4): e0265817. <https://doi.org/10.1371/journal.pone.0265817>

Editor: Tzen-Yuh Chiang, National Cheng Kung University, TAIWAN

Received: March 10, 2021

Accepted: March 8, 2022

Published: April 26, 2022

Copyright: © 2022 Nasar et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Relevant data can be accessed through the following Dryad DOI: <https://doi.org/10.5061/dryad.c866t1g8s> (contains morphological data and script for variant calling and data analyses), and SRA: PRJNA757521 (contains the raw RNA-seq data).

Funding: Incidental materials for genetic analyses funded in part by National Science Foundation (USA) grant DEB 1542387 to MDR.

Abstract

Phaseolus vulgaris, an essential food and source of protein, is cultivated across the world. This study was carried out to investigate the diversity and population structure of 34 *P. vulgaris* landrace accessions collected from the Azad Jammu and Kashmir (AJ&K) regions of Pakistan. The samples were analyzed both morphologically and using genetic variation identified through RNA sequencing. Our results indicated that most genetic variation occurs among local accessions, with little genetic variation occurring between geographical regions. In addition, the accessions fell into two major genetic groups. Morphological analysis revealed that these two genetic groups differ in a number of quantitative traits, including seed length, seed width, and seed weight. One accession, DUD-11, appears to be a mixture of the two major groups genetically as well as morphologically. Among the other accessions, DUD-8, RWK-2, and NGD-1 depicted particularly high seed weight along with higher seed length, seed width, and seed yield per plant. We suggest focusing on these accessions in future breeding programs. More generally, our results provide baseline data that will be useful for crop improvement and effective cultivation practices in Pakistan.

Introduction

The legume family, Leguminosae, is the second most important family among crop plants after Poaceae [1] and contains five *Phaseolus* species that are cultivated worldwide for the commercial production of beans. Common bean, *Phaseolus vulgaris* L. ($2n = 2x = 22$), is the most economically important bean species, having an 85% share in worldwide bean production, followed by *P. lunatus*, *P. coccineus*, *P. acutifolius* and *P. dumosus* [2]. In fact, common bean is the most important agricultural crop after cereals, with an annual global dry yield of more than 8741 hg/ha over 33 million ha [3]. Furthermore, *P. vulgaris* is considered a functional food because it is a rich source of protein, fiber, vitamins and essential minerals [4]. Because

Competing interests: The authors have declared that no competing interests exist.

common bean is such an important crop, breeders are interested in the amount and distribution of genetic and morphological variation in the species.

Previous studies have assessed morpho-agronomic and genetic diversity in common bean and its close relative runner bean (*P. coccineus*). For example, Rosales-Serna et al. [5] carried out genetic analysis of Mexican common bean and runner bean landraces and cultivars by using AFLP markers. Their results showed a broad range of genetic diversity within and between Mexican bean races and that *P. coccineus* cultivars are clearly distinct from *P. vulgaris*. Likewise, Guerra-Garcia et al. [6] reported the *P. coccineus* domestication history and high level of genetic diversity among wild and cultivated populations of Mexico and Spain. In addition, evidence based on morphological [7], biochemical [8, 9], and molecular analyses [10, 11] suggests that there are two distinct gene pools for the common bean, known as Mesoamerican (small-seeded type) and Andean (large-seeded type). These gene pools have diverged from a wild ancestral *P. vulgaris* because of geographical isolation over the last approximately 100,000 years [12]. Common beans have been domesticated from each of these gene pools [13] and cultivars have been dispersed world-wide from both of these centers of origin. There are strong indications that Portuguese, Dutch, English, and/or French travelers bought the common bean with them for trading in the early 16th century when they travelled through Indian sub-continent. However, there is very little known about the dissemination of common beans in the Himalayan region where a rich diversity of beans occurs [14]. Although dispersal outside of Mesoamerica and the Andes region has reduced diversity [15], most of the previous studies demonstrate substantial genetic and morphological diversity in beans, with clusters of similar cultivars that are distinct from other clusters. Many of these studies have focused on Mexican or Andean cultivars [16], while others have focused on either seed banks or cultivars in countries like India and Brazil [17–19].

P. vulgaris is grown extensively in the northern areas of Pakistan, including the Azad Jammu and Kashmir (AJ&K) regions, where the average annual yield is 0.5t/ha and can be as high as 2.5–5t/ha, under the best cultivation practices. Currently, production is unable to meet the national demand [20, 21]. Much of current production is based upon small land holdings with centuries-old customs and cultivation of local landraces [22]. The development of such landraces can lead to genetic divergence between localities, counteracting the reduction of diversity associated with dispersal bottlenecks. However, because there have been no previous investigations of genetic diversity among these land races, the amount and geographic distribution of variation available for crop improvement in Pakistan is currently unknown. There has not been even a single registered variety of common bean reported so far in Pakistan [23–25]. It is thus unknown how much genetic variation exists within the AJ&K region, whether it contains material from both centers of domestication, and particularly if that variation is sufficient for use in crop improvement [26, 27]. Thus, there is a great need to identify high yield common bean accessions that enhance the livelihood of our farmers, to make Pakistan self-sufficient for beans and proves to be profitable in foreign exchange [21].

In a broad sense, the current research aims to assess landrace genotypes that could perform better in low input agriculture systems, to conserve the selected germplasm from possible extinction, and to provide baseline data of selected germplasm for further research purposes. The specific objective of this study is to characterize morphological and genetic variation among Pakistani cultivars and regions of the country. To do this, we address the following questions: (1) How variable are important morphoagronomic traits in Pakistan? (2) What is the heritability of those traits? (3) How much genetic variation is present in Pakistani cultivars? (4) How is the morphological and genetic variation distributed geographically? (5) How does the amount of genetic variation in Pakistan compare to variation present in the *P. vulgaris*

Table 1. *Phaseolus vulgaris* accessions and geographical coordinates of sampling sites in Azad Jammu and Kashmir, Pakistan.

District	Sampling site	Latitude (N)	Longitude (E)	Accessions codes	Species
Neelum	Nagder	34°40'47"	73°56'14"	NGD_13	<i>P. coccineus</i>
				NGD_1	<i>P. vulgaris</i>
	Dodnyal	34°42'08" - 34°40'30"	74°06'25" - 74°05'31"	DUD_1, DUD_2, DUD_4, DUD_8, DUD_11, DUD_12	<i>P. vulgaris</i>
	Kel	34°48'31" - 34°49'27"	74°21'06" - 74°21'07"	KEL_2, KEL_3, KEL_5, KEL_8, KEL_11, KEL_14	<i>P. vulgaris</i>
	Halmat	34°45'09"	74°39'34"	HAL_2, HAL_6	<i>P. vulgaris</i>
Hattian (Leepa valley)	Gai pora	34°18'41" - 34°18'59"	73°51'01"	LPA_1, LPA_3, LPA_4, LPA_5	<i>P. vulgaris</i>
	Nokot	34°18'23"	73°54'31"	LPA_6, LPA_9	<i>P. vulgaris</i>
Muzaffarabad	Machiyara	34°29'58"	73°37'02"	MAC_2, MAC_3, MAC_4	<i>P. vulgaris</i>
Haveli (Farwad Kahuta)	Khurshidabad	33°55'24"	74°10'26"	FK_1, FK_2, FK_3	<i>P. vulgaris</i>
	Kirni	33°56'27"	74°13'19"	FK_4, FK_7, FK_8, FK_9	<i>P. vulgaris</i>
Poonch (Rawalakot)	Chhota gala	33°49'21"	73°48'13"	RWK_1, RWK_2	<i>P. vulgaris</i>
Bagh	Saver	34°01'31"	73°48'57"	BG_1	<i>P. vulgaris</i>

Ranges for Latitude and Longitude given for sampling sites at which accessions differ in these values.

<https://doi.org/10.1371/journal.pone.0265817.t001>

centers of origin in the Americas? These analyses should prove helpful in guiding breeding programs for common bean improvement in Pakistan.

Materials and methods

Thirty-four accessions of *P. vulgaris* and one accession of *P. coccineus* were collected from AJ&K. The *P. vulgaris* accessions were collected from six districts with 1–4 sampling sites in each district and a single *P. coccineus* accession was collected from the Nagdar site in the Neelum district (see [Table 1](#) and [S1 Table](#) for details).

Measurement of morphological characters

Quantitative morphological traits were measured in a field experiment involving 34 *P. vulgaris* accessions and the one *P. coccineus* accession. Seeds were planted on May 1st in 2017 and 2018 in a randomized complete block design with three replicates at a location in Muzaffarabad (latitude: 34°22'10" N, longitude: 73°31'30" E, altitude 1,731 m). The site has an average annual rainfall of 1300 mm and a mean annual temperature of 16 °C. The soil in fields in Muzaffarabad is basic in nature with sufficient organic matter [28]. The field was prepared by deep ploughing, followed by row cultivation, with 80 cm between rows and 40 cm between plants within a row. Plants were allowed to twine up rope supports and weeds were removed periodically.

Thirteen quantitative morphological traits were measured on each of 10 randomly chosen plants per accession at various plant growth stages from seed germination to seed formation following methods recommended by the International Board for Plant Genetic Resources *Phaseolus L.* Descriptor list [29] ([Table 2](#)).

RNA extraction and cDNA library construction

We use RNA-sequencing (RNAseq) of leaf tissue to determine the sequences of genes expressed in leaf tissue and identify single-nucleotide polymorphisms (SNPs) that are subsequently used to characterize patterns of genetic diversity in *P. vulgaris*. RNAseq is commonly used to identify SNPs for characterizing genetic diversity and constructing phylogenies [30–37]. Molecular characterization of genetic variation was conducted on one of the ten

Table 2. List of investigated quantitative traits following International Board for Plant Genetic Resources list (IBPGR, 1982).

Traits	Quantitative traits measurements
Days to flowering	Days to flowering (DTF) were observed from the date of sowing to the date on which 50% of the plants in a plot had set flowers.
Plant height	After maturity, plant height (PH) was measured by using meter scale from first node up to the tip of the plant.
Leaflet length	Terminal leaflet length (LL) of third trifoliate leaf was measured in cm from Pulvinus to the leaf tip for randomly selected plants.
Leaflet width	Leaflet width (LW) of terminal third trifoliate in cm, was measured for randomly selected plants.
Stem girth	At maturity, stem girth (SG) was determined by wrapping a string around plant stem followed by measuring string using graduated meter ruler and obtaining the circumference in cm.
Pod length	Pod length (PL) in cm, was measured from the tip to the peduncle from 10 randomly selected fully expanded immature pods per plant.
Pod width	Pod width (PW) of the fully expanded pods was measured in cm by randomly selected 10 fully expanded immature pods per plant.
Seeds per pod	Total number of seeds per pod (SPP) was counted by taking average number of seeds from randomly selected pods per plant.
Pod Beak length	Pod beak length (PBL) was recorded in cm from end of last locule from randomly selected 10 pods per plant.
Seed length	Seed length (SL) was measured parallel to the hilum in mm through randomly selected ten seeds per plant by using Vernier caliper.
Seed width	Seed width (SW) was measured from hilum opposite side in mm by selecting randomly ten seeds per plant using Vernier caliper
Hundred seeds weight	Hundred dry seeds (HSW) of three randomly drawn samples from each experimental plot were taken, weighed in grams and averaged.
Seed yield per plant	Seed yield of selected plants (SYPP) from each plot was weighed and averaged to single plant basis in grams.

<https://doi.org/10.1371/journal.pone.0265817.t002>

individuals from each of the 35 accessions on which morphological measurements were performed. RNA was extracted from approximately 100mg of leaf tissue from plants at the 2-leaf stage following the protocol of the Spectrum Plant Total RNA Kit (Sigma-Aldrich). After extraction, the concentration, purity and integrity of RNA were examined by using a Nano-drop apparatus (Thermo Fisher Scientific, Waltham, MA, USA) and gel electrophoresis using 1% agarose gels. cDNA libraries were prepared using the KAPA Stranded RNA-Seq library preparation kit (Roche) following the manufacturer's protocol. A separate library was created for each accession. The quality, quantity, and size of the libraries were checked using a Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) and a Bioanalyzer system (Agilent Technologies, Santa Clara, CA, USA). Libraries were sequenced (150 bp paired-end reads) over two lanes on the Illumina HiSeq 4000 platform (Illumina, Inc., San Diego, CA, USA) yielding 627 million paired reads (12–26 million per accession). The sequencing quality was examined using the FastQC tool [38]. Very low-quality bases and reads and reads containing primer or adapter contamination were removed using the software Trimmomatic 0.36 [39]. Specifically, a 4-bp sliding window was used to clip reads that fell below an average PHRED score of 5, and bases with scores below 5 were removed from the start and end of reads. Reads that were less than 25 bp after trimming were discarded. Overall, much less than 1% of reads were discarded. A reference transcriptome was constructed using data from the individual with highest coverage (LPA_9) and the trinity software [40] with the default settings. This resulted in 72,927,950 bases assembled into 74,575 transcripts.

Variant detection. To detect SNPs and indels among the accession, the filtered reads of each accession were aligned to the reference transcripts (75–85% of reads successfully aligned)

and SNPs and indels were called using Samtools mpilup [41]. We filtered all sites (variant and invariant) for least 100 reads and a map quality score of at least 30. This yielded 12,055,420 total sites, of which 269,883 were variant (224,480 SNPs and 45,403 indels; see S1 Fig for histograms of the heterozygosity and minor allele frequencies of these variants).

We also generated a more stringently filtered set of variants. In this case, biallelic variants were filtered for a minimum minor allele frequency of 1% and complete coverage across individuals (i.e., no missing data), and thinned to a single variant per 1kb. This resulted in a total of 15,149 variants.

Statistical analysis

Descriptive statistics were calculated for both the morphological and genetic datasets. To determine whether accessions differed for the morphological traits, we performed a standard analysis of variance (ANOVA) using Statistix 8.1 [42]. Phenotypic (V_P) and genetic (V_G) variance components were calculated using:

$$V_G = \frac{\text{Genotype mean square (GMS)} - \text{Error mean square (EMS)}}{\text{Number of replications (r)}}$$

$$V_P = V_G + V_E$$

Where V_E = Error mean square [43]

Broad-sense heritability was calculated as $H^2 = \frac{V_G}{V_P}$ [44] with the statistical significance of the heritability assessed using an ANOVA with replicate nested within accession. The genetic relationship among accessions was examined, using the stringently filtered set of variants, in three ways. First, we performed a principal component analysis (PCA) the R package SNPRe-relate [45]. Second, we performed a cluster analysis, which groups accessions on the basis of similarity index, and plotted the resulting dendrogram in SNPRe-relate. Finally, we performed a STRUCTURE analysis [46] implemented in StrAuto [47] with 10 replicate runs with 1,000 iterations after 1,000 burn-in iterations for each of 10 potential population clusters ($K = 1-10$). We used Structure Harvester [48] and the ΔK approach [49] to determine the number of population clusters that best fit our data.

To further characterize genetic diversity, we calculated nucleotide diversity (π) both within the 34 samples, as well as within and between two clusters of samples identified by STRUCTURE and PCA. Pairwise π , the proportion of nucleotides by which two sequences from different individuals differ, was calculated using an APL program written by one of the authors (MDR). This program first assigns SNP, S , equal to 0 if the SNP genotype was the same for the two samples, 1 if the two SNPs were homozygous for different alleles, and 0.5 otherwise. Scores were summed to produce S_T . To correct bias caused by missing values, a corrected sum was calculated as $S_C = S_T (1 + F)$, where F is the fraction of SNPs that were missing in at least one sample. π was then calculated as $\pi = S_C/N$, where N is the total number of nucleotides in the transcriptome that passed filtering criteria (12,055,420). The program also calculates average π and its standard error, and performs a bootstrap analysis with 1,000 replicates to estimate confidence intervals.

To determine the extent of geographical trait differentiation, we performed a multiple analysis of variance (MANOVA) in which the factors were Among Districts (random effect), Among Sites Within Districts (nested, random effect) and Within-Sites using PROC GLM in the SAS software package, Version 9.4, for Windows. Variance components were estimated using PROC VARCOMP with the REML analysis. Variance components were converted to

the proportion of variation explained by each effect. Following the MANOVA, univariate ANOVA analyses were performed.

The relationship between trait variation and genetic variation was examined by determining whether traits differed for the two major clusters identified by the analyses described above. Specifically, we performed both MANOVA and univariate ANOVA on trait values in which the main effect was cluster. We excluded the accession DUD_11 from this analysis because it fell outside the two main clusters and because our STRUCTURE analysis suggests that it is a hybrid between the genotypes represented by the two clusters.

Results

Morphological traits

Analysis of Variance indicates that all morphological traits differed among accessions ($P < 0.001$ for all traits in both years; [S2 Table](#)). Broad-sense heritability was high for all traits, ranging from 0.54 to 0.99 ([Table 3](#)).

Comparison of morphological traits across years revealed that for all traits there was a very strong correlation between trait value in 2017 and trait value in 2018 ([S2 Fig](#)). Correlation coefficients ranged from 0.944 to 1.0. Because of these high correlations, in subsequent analyses we used the average trait value across years to simplify the analyses ([S3 Table](#)).

There was little evidence of geographic differentiation in traits. A MANOVA analysis revealed no evidence for significant variation either among districts or among sites within districts. We performed two analyses. In the first, we included both district and site nested within district. In this model, the district effect could not be tested because there were too few error degrees of freedom. By contrast, the site effect was not significant (Wilks Lambda = 0.03326, $F = 0.83$, $df = 65$, 51.202, $P = 0.761$). The second model included only the district effect, which was also not significant (Wilks Lambda = 0.011, $F = 1.26$, $df = 65$, 51.202, $P = 0.198$).

Univariate analyses of individual traits yielded a similar conclusion. Variance components for both district and site effects were generally small and non-significant ([Table 4](#)). The only exception is for pod width, for which there was no significant site effect, and the district effect was significant when there was no site effect in the model. The combined multivariate and univariate analyses thus indicated that most variation for traits is local: it is among accessions within sites.

Table 3. Genetic variance components and broad sense heritability for 13 morphological traits of 34 *Phaseolus vulgaris* accessions.

Traits	Phenotypic variance		Genotypic variance		Heritability (bs)	
	2017	2018	2017	2018	2017	2018
Days to flowering	4202.03	4386.06	4193.03	4383.26	0.93	0.89
Plant height	2.52	2.50	1.71	1.67	0.99	0.99
Leaflet length	2.03	2.02	1.10	1.09	0.68	0.67
Leaflet width	2.18	2.17	1.49	1.49	0.54	0.54
Stem girth	0.06	0.05	0.04	0.03	0.86	0.88
Pod length	0.04	0.04	0.04	0.03	0.69	0.69
Pod width	0.33	0.26	0.23	0.24	0.67	0.67
Pod beak length	153.85	158.25	147.82	152.41	0.87	0.85
Number seeds per pod	3.56	3.56	3.08	3.08	0.72	0.93
Seed length	1.01	1.01	0.83	0.83	0.87	0.87
Seed width	0.32	0.29	0.27	0.26	0.82	0.82
Hundred Seed weight	44.24	46.36	43.46	44.04	0.96	0.96
Seed yield per plant	18.01	21.11	16.79	18.85	0.98	0.95

<https://doi.org/10.1371/journal.pone.0265817.t003>

Table 4. Variance components (proportion of variance explained) for different levels of geographic variation for 13 morphological traits.

Traits	Districts	P1	P2	Site	P3	Within site
Days to flowering	0.000	0.470	0.587	0.000	0.474	1.0
Plant height	0.000	0.727	0.325	0.295	0.030	0.705
Leaflet length	0.000	0.858	0.973	0.000	0.488	1.000
Leaflet width	0.000	0.956	0.969	0.010	0.129	0.990
Stem girth	0.092	0.307	0.049	0.325	0.063	0.583
Pod length	0.000	0.492	0.781	0.000	0.960	1.000
Pod width	0.000	0.235	0.430	0.000	0.789	1.000
Pod beak length	0.434	0.020	0.0034*	0.000	0.357	0.566
Number seeds per pod	0.000	0.388	0.480	0.000	0.443	1.000
Seed length	0.081	0.116	0.218	0.000	0.822	0.919
Seed width	0.093	0.154	0.288	0.000	0.837	0.907
Hundred Seed weight	0.145	0.059	0.174	0.000	0.839	0.855
Seed yield per plant	0.000	0.609	0.552	0.003	0.572	0.997

P1: P value for test of District effect over synthesized denominator incorporation Site and Within-site effects. P2: P value for model with no Site effect included (appropriate where site effect not significant). P3: P value for Site effect in full model.

*Significant after sequential Bonferroni correction for 13 traits (i.e., at $P = 0.05/13 = 0.00385$).

<https://doi.org/10.1371/journal.pone.0265817.t004>

Patterns of genetic variation

A PCA on the *P. vulgaris* and *P. coccineus* accessions revealed that the first two principal components accounted for 75% of the genetic variation. Principal component one separates *P. coccineus* from *P. vulgaris* and principal component two separates *P. vulgaris* into two major groups (Fig 1A; groups are color coded). A second PCA performed on only the 34 *P. vulgaris* accessions also indicated that *P. vulgaris* is divided into two major genetic groups, this time along the first principal component, which accounts for 48% of the variation. In both analyses, accession DUD_11 falls between the two *P. vulgaris* groups, suggesting it may be a hybrid between the two groups.

A cluster analysis revealed groupings that are consistent with those obtained from the PCA (Fig 1C). The accessions contained within these two groups are identical with those in the two groups identified by PCA. Additionally, accession DUD_11 is an outgroup to group A and to group B, consistent with it being between the two clusters in the PCA analysis.

Finally, a STRUCTURE analysis also yielded a pattern consistent with the previous two analyses. The delta K vs. K plot (S3 Fig) shows a single peak at $K = 2$. Examination of the STRUCTURE plots for $K = 2$ and $K = 3$ revealed that they are very similar, except that at $K = 3$ the *P. coccineus* individual constitutes a separate group. In both cases, all but one of the *P. vulgaris* accessions belongs to one of two major genetic groups (Fig 1D and 1E red vs. blue). The accessions in these two groups correspond precisely to the accessions in the two groups identified by PCA and cluster analysis. Once again, DUD_11 is the exception in that it appeared to contain genetic elements from both groups, consistent with it being a hybrid.

In accordance with the previous analyses, nucleotide diversity (π) was approximately four times greater between the red and blue clusters than within either cluster (Table 5). Although nucleotide diversity was significantly higher in the red cluster ($P < 0.001$, 1000 bootstrap replicates), the difference was not large.

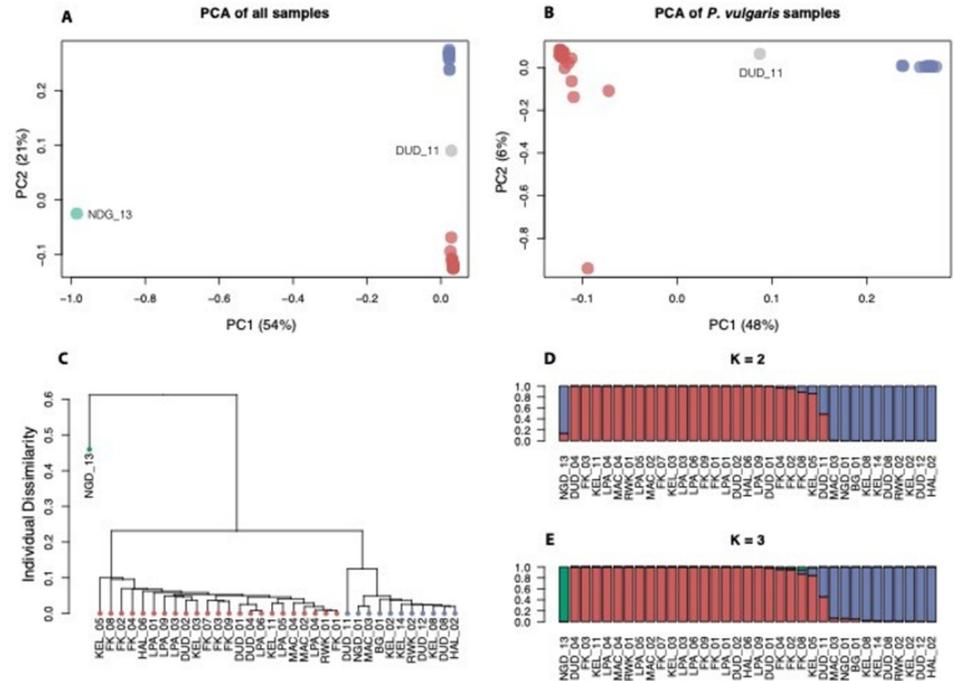


Fig 1. Genetic analyses of 34 *P. vulgaris* accessions and one *P. coccineus* accession (NGD_13). A. Principal component analyses on all individuals. B. Principal component analyses on the *P. vulgaris* individuals. C. Cluster analysis on all individuals. D. STRUCTURE plot for K = 2. E. STRUCTURE plot for K = 3. 218. Red and blue colors correspond to inferred Mesoamerican and Andean accessions, respectively.

<https://doi.org/10.1371/journal.pone.0265817.g001>

Relationship between morphological and genetic variation

We examined whether the two genetic clusters identified in the previous section are differentiated with respect to the morphological traits by performing a MANOVA and univariate ANOVAs in which the main effect was genetic cluster (the purported hybrid accession DUD_11 was excluded in from this analysis). In the multivariate analysis, the cluster effect was significant (Wilks Lambda = 0.3656, F = 2.54, df = 13, 19, P = 0.032).

The univariate analyses indicated that this differentiation was nominally significant for several traits: leaflet length, leaflet width, pod length, pod width, seed width, seed weight, and seeds per pod (Table 6, S4 Table, and S4 Fig). Each of these traits is larger in the blue cluster, suggesting that the blue cluster corresponds to large-seeded Andean accessions and the red cluster corresponds to small-seeded Mesoamerican accessions. On average, the between-cluster proportion of total variation for these traits is 0.393, indicating a not insubstantial divergence between clusters. For the remaining traits, the proportion of variation between clusters averages only 0.0328, with the proportion within clusters being 0.9672. When a sequential Bonferroni test was applied, leaflet length, pod width, seed width, and seed weight remained

Table 5. Nucleotide diversity within and between red and blue clusters.

	Mean π	π Standard Error	π 95% CI (bootstrap)
All 34 Accessions	0.00289	0.00006	(0.00275, 0.00303)
Between Red and Blue Clusters	0.00484	0.000074	(0.00481, 0.00487)
Within Red Cluster	0.00137	0.000027	(0.00132, 0.00142)
Within Blue Cluster	0.00110	0.000034	(0.00103, 0.00117)

<https://doi.org/10.1371/journal.pone.0265817.t005>

Table 6. Variance components (proportion of variance explained for the two different genetic clusters for 13 morphological traits).

Traits	Between-site Variance Component	Within site Variance Component	P
Days to flowering	0.015	0.985	.280
Plant height	0.000	1.000	.928
Leaflet length	0.458	0.542	.0015*
Leaflet width	0.309	0.691	.0136
Stem girth	0.000	1.000	.554
Pod length	0.229	0.771	.0347
Pod width	0.405	0.595	.0036*
Pod beak length	0.182	0.818	.057
Number seeds per pod	0.000	1.000	.686
Seed length	0.280	0.720	.0192
Seed width	0.433	0.567	.0023*
Hundred Seed weight	0.636	0.364	< .001*
Seed yield per plant	0.000	1.000	.450

P: P value for test of Cluster effect in ANOVA.

*Significant after sequential Bonferroni correction for 13 traits.

<https://doi.org/10.1371/journal.pone.0265817.t006>

significant at an overall level of 0.05, with an average between-cluster proportion of variance equal to 0.483.

Discussion

In Pakistan, the production of common bean has been centered on local landrace accessions cultivated by small-scale farmers mostly under low-input agriculture [22]. The traditional *P. vulgaris* landraces have been persistent in crop systems despite the import of seeds from other countries to meet local bean demand, presumably because of local cultural values and the inherent quality of local common bean genotypes [23]. The local accessions exhibit considerable morphological diversity, which reflects rich germplasm diversity owing to geographic and microclimatic variation in bean producing areas in the mountainous region of Kashmir [14]. In the current study, *P. vulgaris* accessions were characterized both morphologically and genetically to examine geographical patterns of variation.

Our survey of diversity among *P. vulgaris* bean cultivars in AJ&K, Pakistan has revealed two major patterns. The first is that there is little large-scale geographic differentiation, either genetically or in trait characteristics. Instead, most variation occurs within sampling sites. A similar lack of geographic differentiation has been reported in other countries [50, 51]. One likely explanation of this pattern is the tradition of sowing mixed seeds. Subsistence farmers tend to keep seed mixtures intact and impose little selection on their landraces, resulting in high local diversity [52]. In addition, variation within localities may be due to temporal heterogeneity associated with short-term climatic variation, with different varieties performing better in different seasons [53]. Finally, germplasm exchange among geographical regions may also contribute to this pattern [51, 54].

A second pattern is that most of the genetic variation in traits is between two major genetic clusters with similar geographic distributions, as has been found in previous studies in other countries [54–56]. Identical clustering was revealed by the STRUCTURE analysis, by cluster analysis and by Principal component analysis. The smaller cluster (Cluster A, blue) was comprised of 10 large-seeded accessions, based on mean HSW > 40g/100 seeds. These accessions also tended to have large leaves and pods. The 23 accessions in the larger cluster (Cluster B,

red) had smaller seeds (< 30g/100 seeds). These two clusters presumably correspond to different geographic origin. Mesoamerican varieties tend to have small to medium sized seeds, while Andean varieties tend to have large seeds [57–59]. Although it would be useful to confirm this inference by comparing sequence similarity, it would thus seem that the large-seeded cultivars in our study are likely derived from the Andes, while those with smaller seeds are likely derived from Mesoamerica. Assuming that the relative proportions of large- and small-seeded cultivars in our study are representative of the proportions in AJ&K, Pakistan as a whole, there are more accessions belonging to Mesoamerican gene pool than Andean gene pool. The apparent preference for small-seeded accessions in the AJ&K region may be attributed to the desirable traits such as taste, seed size, growth habit, and seed colour. Such preferences have also been recorded in several other parts of the world including China [52, 60, 61], India [14, 62], Brazil [11] and Italy [63]. This distribution difference might be related to the germplasm introduction time, germplasm adaptation abilities, ecological types and consumer preferences.

Genetic diversity in the Mesoamerican samples is considerably less than that reported for land races from this region. Our estimate of π for Pakistani Mesoamerican derivatives (0.00137) is less than $\frac{1}{2}$ to $\frac{1}{7}$ that of Mesoamerican land races (0.0099, Mamidi et al. [64]; 0.0030, Bitocchi et al. [65]). Our estimate of π for Pakistani Andean derivatives (0.00110) is between estimates for Andean land races (0.0066, Mamidi et al. [64]; 0.00074, Bitocchi et al. [65]), but is only $\frac{1}{6}$ of the higher estimate. While a small number of genes were used in these previous studies to estimate π in Andean accessions (13 and 5 for Mamidi et al. [64] and Bitocchi et al. [65], respectively) we believe the reduced diversity of Mesoamerican derivatives in Pakistan is likely real, and may also be true for Andean derivatives. This reduction suggests severe bottlenecks occurred during dispersal of the common bean to Pakistan. Additionally, it suggests genetic variation for crop improvement, especially of Mesoamerican derivatives, in Pakistan may be severely limited and that breeders should consider introducing genetic material from other regions.

Two accessions in this study fall outside the two major genetic clusters. One is NGD-13, which is not surprising because this accession represented a different species i.e., *P. coccineus*. Of more interest is accession DUD_11 (S5 Table), which lies between the two main groups in the PCA analysis and in the dendrogram. The STRUCTURE analysis demonstrated that this accession consists of a mixture of genotypes characteristic of the two main clusters, suggesting that it represents a natural hybrid between one accession from each cluster. Unfortunately, because all trait values lie within the range of values for both clusters (S5 Table), traits of this accession are not informative about hybridization. This type of hybridization is not unexpected, given that accessions from both clusters frequently grow at the same locations. This type of hybridization has been found in *P. vulgaris* from other regions [66] as well as in CIAT core collections [67]. Such hybrid accessions may be of interest to breeders because they constitute germplasm bridging the Andean and Mesoamerican gene pools and thus may be helpful in improving Andean and Mesoamerican varieties [16].

Our study has revealed an interesting trend regarding growth form, as most of the accessions we examined were indeterminate climbers growing to the taller heights. This trait may be beneficial in Pakistan because beans tend to be grown as intercrops, often with maize, with the bean plants climbing the maize stalks. A tall stature may thus be an eco-physiological adaptation for attaining maximum light exposure and increasing their photosynthetic efficiency. The indeterminate climber growth habit is reported to be the best adapted for mountainous areas with cooler and wet climates [68] such as those found in the AJ&K region, Pakistan.

A comparative analysis of the literature reveals that our leaf size values are considerably higher than those reported by Bode et al. [69] and Loko et al. [70]. The difference in leaf area is significantly correlated with the geographical distribution of common bean genotypes as an

adaptation in cooler agroclimatic zones. Consequently, cool temperature of the temperate mountainous environment facilitated the leaf expansion. Duncan and Hesketh [71] also reported greater leaf area in highland cold temperatures adapted maize genotypes as compared to lowland-adapted genotypes. It was also observed that large-seeded genotypes had large surface area. It has been reported that Andean lines exhibit larger leaves attributed to a selection pressure for larger leaves caused by cooler temperatures, which inversely affects the leaf thickness [72]. The lower leaf thickness also facilitates the tall climber genotypes to bear more leaves by reducing the leaf weight.

Our results also indicate relatively long Days to flowering in the investigated genotypes averaging as 65 days. Days to flowering is an important parameter which reflects the physiological response of plant growth mechanisms influenced by the climatic factors, including temperature and precipitation [70]. This trait may have evolved via changes in photoperiodism sensitivity and adaptation to the Cooler temperature of the Himalayan temperate region in which common bean landraces have been growing [72].

Conclusion

The present study revealed considerable morphological and genetic variation among common bean accessions from Azad Jammu and Kashmir, Pakistan. There is little geographic differentiation either genetically or morphologically and much variation occurs locally within sites. Accessions were divided into two genetic groups that differed in seed length and width, hundred seed weight, and seed yield, presumably reflecting Andean and Mesoamerican sources. One accession lies outside the two genetics groups and is more likely a hybrid between the two groups. The depauperate variation in Pakistani derivatives from Mesoamerica, and possibly from the Andes, suggests that breeding programs should include accessions from outside Pakistan.

Supporting information

S1 Fig. Histograms of heterozygosity and minor allele frequency for the variants we detected.

(DOCX)

S2 Fig. Correlations of means across years for the 13 morphological traits. Each point represents one accession.

(DOCX)

S3 Fig. Evanno plot for determining the number of distinct genetic groups in the STRUCTURE analysis of 34 *P. vulgaris* accessions.

(DOCX)

S4 Fig. Box plots for morphological traits comparing accessions in the two genetic clusters. Colors of clusters correspond to colors of clusters in Fig 1.

(DOCX)

S1 Table. Inventory of Phaseolus accessions and geographical characters of sampling sites in AJK.

(DOCX)

S2 Table. ANOVA for testing differences in morphological traits among accessions in years 2017 and 2018 for 34 Phaseolus vulgaris accessions.

(DOCX)

S3 Table. Mean and Standard error of morphological trait values. Means are averages across 2017 and 2018 for 34 *Phaseolus vulgaris* accessions.

(DOCX)

S4 Table. Statistics for morphological traits by genetic cluster for 34 *Phaseolus vulgaris* accessions. A. Statistics for Blue cluster. B. Statistics for Red cluster.

(DOCX)

S5 Table. Trait values for accession DUD-11 compared to ranges of values for Red (Meso-american) and Blue (Andean) clusters. Low: lowest value of trait among accessions within cluster. High: Highest value of trait among accession within clusters.

(DOCX)

Acknowledgments

HEC Pakistan for provided platforms for international recognition of a Pakistani student to upgrade the standard of local educational research to meet international requirements and goals.

Author Contributions

Conceptualization: Sidra Nasar, Kate Ostevik, Ghulam Murtaza, Mark D. Rausher.

Data curation: Sidra Nasar.

Formal analysis: Sidra Nasar, Kate Ostevik, Mark D. Rausher.

Funding acquisition: Mark D. Rausher.

Investigation: Sidra Nasar.

Methodology: Sidra Nasar, Kate Ostevik, Ghulam Murtaza, Mark D. Rausher.

Project administration: Sidra Nasar, Ghulam Murtaza, Mark D. Rausher.

Supervision: Kate Ostevik, Ghulam Murtaza, Mark D. Rausher.

Visualization: Sidra Nasar.

Writing – original draft: Sidra Nasar, Kate Ostevik, Mark D. Rausher.

Writing – review & editing: Sidra Nasar, Kate Ostevik, Mark D. Rausher.

References

1. Lewis G, Schrire B, Mackinder, Lock B M. Legumes of the World, Royal Botanic Gardens, Kew. London: Cambridge University Press. 2005.
2. Mir RR, Choudhary N, Bawa V, Jan S, Singh B, Ashraf Bhat M, et al. Allelic Diversity, Structural Analysis and Genome-wide Association Study (GWAS) for Yield and Related Traits using Unexplored Common bean (*Phaseolus vulgaris* L.) Germplasm from Western Himalayas. *Frontiers in genetics*. 2020; 11: 1–17. <https://doi.org/10.3389/fgene.2020.00001> PMID: 32117431
3. Food and Agriculture Organization of the United Nations (FAO). (2019). The food and agricultural commodities. [Online]. Available: <http://faostat3.fao.org/home/E>
4. Celmeli T, Sari H, Canci H, Sari D, Adak A, Eker T, et al. The nutritional content of common bean (*Phaseolus vulgaris* L.) landraces in comparison to modern varieties. *Agronomy*. 2018; 8(9):166–175. <https://doi.org/10.3390/agronomy8090166>
5. Rosales-Serna R, Hernandez-Delgado S, Gonzalez-Paz M, Acosta-Gallegos JA, Mayek-Perez N. Genetic relationships and diversity revealed by AFLP markers in Mexican common bean bred cultivars. *Crop Science*. 2005; 45(5):1951–1957. <https://doi.org/10.2135/cropsci2004.0582>

6. Guerra-Garcia A, Suarez-Atilano M, Mastretta-Yanes A, Delgado-Salinas A, Pinero D. Domestication genomics of the open-pollinated scarlet runner bean (*Phaseolus coccineus* L.). *Frontiers in Plant Sciences*. 2017; 8:1891. <https://doi.org/10.3389/fpls.2017.01891> PMID: 29187858
7. Meza N, Rosas JC, Martin JP, Ortiz JM. Biodiversity of common bean (*Phaseolus vulgaris* L.) in Honduras, evidenced by morphological characterization. *Genetic Resource and Crop Evolution*. 2013; 60(4):1329–1336. <https://doi.org/10.1007/s10722-012-9922-7>
8. Koenig R, Gepts P. Allozyme diversity in wild *Phaseolus vulgaris*: further evidence for two major centers of genetic diversity. *Theoretical and Applied Genetics*. 1989; 78(6):809–817. <https://doi.org/10.1007/BF00266663> PMID: 24226011
9. Nogueira AF, Moda-Cirino V, Delfini J, Brandao LA, Mian S, Constantino LV, et al. Morpho-agronomic, biochemical and molecular analysis of genetic diversity in the Mesoamerican common bean panel. *PloS One*. 2021; 16(4):e0249858. <https://doi.org/10.1371/journal.pone.0249858> PMID: 33886606
10. Lioi L, Nuzzi A, Campion B, Piergiovanni AR. Assessment of genetic variation in common bean (*Phaseolus vulgaris* L.) from Nebrodi mountains (Sicily, Italy). *Genetic Resource and Crop Evolution*. 2012; 59(3):455–464. <https://doi.org/10.1007/s10722-011-9696-3>
11. Delfini J, Moda-Cirino V, dos Santos Neto J, Ruas PM, Santana GC, Gepts P, et al. Population structure, genetic diversity and genomic selection signatures among a Brazilian common bean germplasm. *Scientific Reports*. 2021; 11(1):1–12. <https://doi.org/10.1038/s41598-020-79139-8> PMID: 33414495
12. Bitocchi E, Nanni L, Bellucci E, Rossi M, Giardini A, Zeuliet al. Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data. *Proceedings of the National Academy of Sciences*. 2012; 109(14):788–796. <https://doi.org/10.1073/pnas.1108973109> PMID: 22393017
13. Rendon-Anaya M, Montero-Vargas JM, Saburido-Alvarez S, Vlasova A, Capella-Gutierrez S, Ordaz-Ortiz JJ. Genomic history of the origin and domestication of common bean unveils its closest sister species. *Genome biology*. 2017; 18(1): 1–17. <https://doi.org/10.1186/s13059-016-1139-1> PMID: 28077169
14. Banoo A, Nabi A, Rasool RS, Shah MD, Ahmad M, Sofi PA, et al. North-western Himalayan common beans: population structure and mapping of quantitative anthracnose resistance through genome wide association study. *Frontiers in Plant Science*. 2020; 11:1–18. <https://doi.org/10.3389/fpls.2020.00001> PMID: 32117356
15. De Luca D, Cennamo P, Del Guacchio E, Di Novella R, Caputo P. Conservation and genetic characterization of common bean landraces from Cilento region (southern Italy): high differentiation in spite of low genetic diversity. *Genetica*. 2018; 146(1):29–44. <https://doi.org/10.1007/s10709-017-9994-6> PMID: 29030763
16. Kwak M and Gepts P. Structure of genetic diversity in the two major gene pools of common bean (*Phaseolus vulgaris* L., Fabaceae *Theoretical and Applied Genetics*. 2009; 118(5):979–992. <https://doi.org/10.1007/s00122-008-0955-4> PMID: 19130029
17. Shah AH, Khaliq I, Bashir G, Jannat S. Genetic Diversity in Common Bean (*Phaseolus vulgaris* L.) Germplasm to Establish a Breeding Program in the Mountainous Region of Jammu and Kashmir. *Indian Journal of Agricultural Research*. 2020; 54(5): 629–634.
18. Cabral PDS, de Souza LC, da Costa GF, Silva FHL, Soares TCB. Investigation of the genetic diversity of common bean (*Phaseolus vulgaris*) cultivars using molecular markers. *Genetics and Molecular Research*. 2018; 17(4):1–11. <https://doi.org/10.4238/gmr18106>
19. Venkataravanappa V, Sonavane P, Reddy CL, Reddy MK. Characterization of novel French bean leaf curl virus and betasatellite associated with leaf curl disease of French bean (*Phaseolus vulgaris* L.). *Journal of General Plant Pathology*. 2021; 87(2): 94–100. <https://doi.org/10.1007/s10327-020-00979-3>
20. PARC (Pakistan Agriculture Research Council). (2020). National Coordinated Pulses Programme, news on the website of <http://www.parc.gov.pk/>
21. Ali A, Khan SA, Ali S, Ali N, Khan SM, Hussain I, et al. Morphological and biochemical characterization of locally available kidney beans. *Pure and Applied Biology*. 2020; 9(1):528–537. <https://doi.org/10.19045/bspab.2020.90058>
22. Khaliq A, Iqbal MA, Zafar M, Gulzar A. Appraising economic dimension of maize production under coherent fertilization in Azad Kashmir, Pakistan. *Custos E Agronegocio On Line*. 2019; 15(2):243–253.
23. Sher K, Nisar M, Subhan M. PCR based genetic diversity in common bean (*Phaseolus vulgaris*) landraces collected from Malakand, Division, Khyber Pakhtunkhwa, Pakistan. *Sarhad Journal of Agriculture*, 2019; 35(2): 579–586. <http://dx.doi.org/10.17582/journal.sja/2019/35.2.579.586>
24. Asim M. Evaluation of common bean germplasm collected from the neglected pockets of northwest Pakistan at Kalam (Swat). *Pakistan Journal of Botany*. 2011; 43(1): 213–219.

25. Nawaz I, Farhatullah FM, Ali A, Ali GM. Primary evaluation of seed characteristics of common bean landraces collected from Himalaya region of Pakistan. *Sarhad Journal of Agriculture*. 2020; 36(1): 33–41. <http://dx.doi.org/10.17582/journal.sja/2020/36.1.33.41>
26. Jannat S, Shah AH, Shah KN, Kabir S, Ghafoor A. Genetic and nutritional profiling of common bean (*Phaseolus vulgaris* L) germplasm from Azad Jammu and Kashmir and exotic accessions. *The Journal of Animal & Plant Sciences*. 2019; 29(1): 205–214.
27. Nasar S and Murtaza G. Morphoagronomic Characterization and Genetic Variability Analysis of Common bean Landraces in Kashmir Himalayas: Implications for Yield Enhancement. *Pakistan Journal of Botany*. 2022; 54(2): 497–506. [https://doi.org/10.30848/Pjb2022-2\(3\)](https://doi.org/10.30848/Pjb2022-2(3))
28. Abbasi MA and Butt MB. Soil fertility status of cultivated lands in different agro-ecological zones of Azad Jammu and Kashmir, Pakistan. *Journal of Applied Agriculture and Biotechnology*. 2018; 2(2):21–27.
29. IBPGR. (1982). International Board for Plant Genetic Resources. Descriptors List for *Phaseolus vulgaris*, Secretariat Rome, Italy.
30. Wang S, Wong D, Forrest K, Allen A, Chao S, et al. Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnology Journal*. 2014; 12:787–796. <https://doi.org/10.1111/pbi.12183> PMID: 24646323
31. Ophir R, Sherman A, Rubinstein M, Eshed R, Schwager MS, Harrel-Beja R, et al. Single-nucleotide polymorphism markers from de-novo assembly of the pomegranate transcriptome reveal germplasm genetic diversity. *PLOS One*. 2014; 9: e88998. <https://doi.org/10.1371/journal.pone.0088998> PMID: 24558460
32. Romiguier J, Gayral P, Ballenghien M, Bernard A, Cahais V, Chenuil A, et al. Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature*. 2014; 515: 261–263. www.nature.com/doi/10.1038/nature13685 PMID: 25141177
33. Rong J, Lammers Y, Strasburg JL, Schidlo NS, Ariyurek Y., de Jong TJ, et al. New insights into domestication of carrot from root transcriptome analyses. *BMC Genomics*. 2014; 15: 895. <http://www.biomedcentral.com/1471-2164/15/895> <https://doi.org/10.1186/1471-2164-15-895> PMID: 25311557
34. Kuhn DN, Dillon NL, Innes DJ, Wu L-S, Mockaitis K. Development of single nucleotide polymorphism (SNP) markers from the mango (*Mangifera indica*) transcriptome for mapping and estimation of genetic diversity. *Acta Hort.* 1111. ISHS. 2016; 2016: 315–321. <https://doi.org/10.17660/ActaHortic.2016.1111.45>
35. Wang K, Hong W, Jiao H, Zhao H. Transcriptome sequencing and phylogenetic analysis of four species of luminescent beetles. *Scientific Reports*. 2017; 7: 1814. <https://doi.org/10.1038/s41598-017-01835-9> PMID: 28500331
36. Ge Y, Zhang T, Wu B, Tan L, Ma F, Zou M, et al. Genome-wide assessment of Avocado germplasm determined from specific length amplified fragment sequencing and transcriptomes: Population structure, genetic diversity, identification, and application of race-specific markers. *Genes*. 2019; 10: 215. <https://doi.org/10.3390/genes10030215> PMID: 30871275
37. Leebens-Mack JH, Barker MS, Carpenter EJ, Deyholos MK, Gitzendanner MA, Graham SW, et al. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature*. 2019; 574: 679–685. <https://doi.org/10.1038/s41586-019-1693-2> PMID: 31645766
38. Leggett RM, Ramirez-Gonzalez RH, Clavijo B, Waite D, Davey RP. Sequencing quality assessment tools to enable data-driven informatics for high throughput genomics. *Frontiers in Genetics*. 2013; 4: 288. <https://doi.org/10.3389/fgene.2013.00288> PMID: 24381581
39. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014; 30(15):2114–2120. <https://doi.org/10.1093/bioinformatics/btu170> PMID: 24695404
40. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nature Biotechnology*. 2011; 29 (7):644–52. <https://doi.org/10.1038/nbt.1883> PMID: 21572440
41. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and samtools. *Bioinformatics*. 2009; 25(16):2078–2079. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943
42. McGraw-Hill C. 2008. Statistix 8.1 (Analytical Software, Tallahassee, Florida). Maurice/Thomas text.
43. Burton GW and Devane DE. Estimating heritability in tall fescue (*Festuca arundinacea*) from replicated clonal material 1. *Agronomy Journal*. 1953; 45(10):478–481. <https://doi.org/10.2134/agronj1953.00021962004500100005x>
44. Lush JL. Intra-sire correlations or regressions of offspring on dam as a method of estimating heritability of characteristics. *Journal of Animal Science*. 1940; 1940(1):293–301.

45. Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*. 2012; 28(24):3326–3328. <https://doi.org/10.1093/bioinformatics/bts606> PMID: 23060615
46. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*, 2000; 155(2):945–959. <https://doi.org/10.1093/genetics/155.2.945> PMID: 10835412
47. Chhatre VE and Emerson KJ. StrAuto: automation and parallelization of STRUCTURE analysis. *BMC bioinformatics*. 2017; 18(1):1–5. <https://doi.org/10.1186/s12859-016-1414-x> PMID: 28049414
48. Earl DA and Vonholdt BM. Structure Harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*. 2012; 4:359–361. <https://doi.org/10.1007/s12686-011-9548-7>
49. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*. 2005; 14(8):2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x> PMID: 15969739
50. Nadeem MA, Habyarimana E, Ciftci V, Nawaz MA, Karakoy T, Comertpay G, et al. Characterization of genetic diversity in Turkish common bean gene pool using phenotypic and whole-genome DArTseq-generated silicoDArT marker information. *PLoS One*. 2018; 13(10):e0205363. <https://doi.org/10.1371/journal.pone.0205363> PMID: 30308006
51. Gyang PJ, Nyaboga EN, Muge EK. Molecular characterization of common bean (*Phaseolus vulgaris* L.) genotypes using microsatellite markers. *Journal of Advances in Biology & Biotechnology*. 2017; 13(2):1–15. <https://doi.org/10.1007/s40011-019-01101-0>
52. Zhang X, Blair MW, Wang S. Genetic diversity of Chinese common bean (*Phaseolus vulgaris* L.) landraces assessed with simple sequence repeat markers. *Theoretical and Applied Genetics*. 2008; 117(4):629–640. <https://doi.org/10.1007/s00122-008-0807-2> PMID: 18548226
53. Freyre R, Rios R, Guzman L, Deboucq DG, Gepts P. Ecogeographic distribution of *Phaseolus* spp. (Fabaceae) in Bolivia. *Economic Botany*. 1996; 50(2):195–215. <https://doi.org/10.1007/BF02861451>
54. Buah S, Buruchara R, Okori P. Molecular characterisation of common bean (*Phaseolus vulgaris* L.) accessions from Southwestern Uganda reveal high levels of genetic diversity. *Genetic Resources and Crop Evolution*. 2017; 64(8):1985–1998. <https://doi.org/10.1007/s10722-017-0490-8>
55. Nemli S, Aşciogul TK, Ateş D, Eşiyok D, Tanyolac MB. Diversity and genetic analysis through DArTseq in common bean (*Phaseolus vulgaris* L.) germplasm from Turkey. *Turkish Journal of Agriculture and Forestry*. 2017; 41(5):389–404.
56. Valentini G, Goncalves-Vidigal MC, Elias JCF, Moiana LD, Mindo NNA. Population structure and genetic diversity of common bean accessions from Brazil. *Plant Molecular Biology Reporter*. 2018; 36(5):897–906. <https://doi.org/10.1038/s41598-021-82437-4>
57. Evans AM. Plant architecture and physiological efficiency in the field bean. In: D. Wall editor. *Potentials of Field Bean and other Food Legumes in Latin America*. Cali, Colombia: CIAT; 1973. pp. 279–284.
58. Evans AM. *Structure, variation, evolution, and classification in Phaseolus*. Kew, Richmond, Surrey, UK Royal Botanic Gardens; 1980.
59. Lynch M and Milligan BG. Analysis of population genetic structure with RAPD markers. *Molecular Ecology*. 1994; 3(2):91–99. <https://doi.org/10.1111/j.1365-294x.1994.tb00109.x> PMID: 8019690
60. Lei L, Wang L, Wang S, Wu J. Marker-trait association analysis of seed traits in accessions of common bean (*Phaseolus vulgaris* L.) in China. *Frontiers in genetics*, 2020; 11: 698. <https://doi.org/10.3389/fgene.2020.00698> PMID: 32714377
61. Wu J, Wang L, Fu J, Chen J, Wei S, Zhang S, et al. Resequencing of 683 common bean genotypes identifies yield component trait associations across a north–south cline. *Nature Genetics*. 2020; 52(1): 118–125. <https://doi.org/10.1038/s41588-019-0546-0> PMID: 31873299
62. Rana JC, Sharma TR, Tyagi RK, Chahota RK, Gautam NK, Singh M, et al. Characterisation of 4274 accessions of common bean (*Phaseolus vulgaris* L.) germplasm conserved in the Indian gene bank for phenological, morphological and agricultural traits. *Euphytica*. 2015; 205(2): 441–457. <https://doi.org/10.1007/s10681-015-1406-3>
63. Ilaria M, Alessandra B, Maurizio M, Pietro C, Giovanni D. Characterization of some Italian common bean (*Phaseolus vulgaris* L.) landraces by RAPD, semi-random and ISSR molecular markers. *Genetic Resource and Crop Evolution*. 2007; 54:175–188. <https://doi.org/10.1007/s10722-005-3133-4>
64. Mamidi S, Rossi M, Annam D, Moghaddam S, Lee R, Papa R, et al. Investigation of the domestication of common bean (*Phaseolus vulgaris*) using multilocus sequence data. *Functional Plant Biology*. 2011; 38(12): 953–967. <https://doi.org/10.1071/FP11124> PMID: 32480954
65. Bitocchi E, Bellucci E, Giardini A, Rau D, Rodriguez M, Biagetti E, et al. Molecular analysis of the parallel domestication of the common bean (*Phaseolus vulgaris*) in Mesoamerica and the Andes. *New Phytologist*. 2013; 197(1): 300–313. <https://doi.org/10.1111/j.1469-8137.2012.04377.x>

66. Lobaton JD, Miller T, Gil J, Ariza D, De la Hoz JF, Soler A, et al. Resequencing of common bean identifies regions of inter-gene pool introgression and provides comprehensive resources for molecular breeding. *The Plant Genome*. 2018; 11(2):1–21. <https://doi.org/10.3835/plantgenome2017.08.0068> PMID: 30025029
67. Nkhata W, Shimelis H, Melis R, Chirwa R, Mzengeza T, Mathew I, et al. Population structure and genetic diversity analyses of common bean germplasm collections of East and Southern Africa using morphological traits and high-density SNP markers. *Plos One*. 2020; 15(12): e0243238. <https://doi.org/10.1371/journal.pone.0243238> PMID: 33338076
68. Zizumbo-Villarreal D, Colunga-GarciaMarin P, de la Cruz, EP Delgado-Valerio P, Gepts P. Population structure and evolutionary dynamics of wild–weedy–domesticated complexes of common bean in a Mesoamerican region. *Crop Science*. 2005; 45(3):1073–1083. <https://doi.org/10.2135/cropsci2004.034>
69. Bode D, Elezi F, Gixhari B. Morphological characterisation and interrelationships among descriptors in *Phaseolus vulgaris* accessions. *Agriculture & Forestry*. 2013; 59(2):175–185.
70. Loko LEY, Orobisiyi A, Adjatin A, Akpo J, Toffa J, Djedatin G, et al. Morphological characterization of common bean (*Phaseolus vulgaris* L.) landraces of Central region of Benin Republic. *Journal of Plant Breeding and Crop Sciences*. 2018; 10(11):304–318. <https://doi.org/10.5897/JPBCS2018.0766>
71. Duncan WG and Hesketh JD. Net photosynthetic rates, relative leaf growth rates, and leaf numbers of 22 races of maize grown at eight temperatures 1. *Crop Science*. 1968; 8(6):670–674. <https://doi.org/10.2135/cropsci1968.0011183X000800060009x>
72. Sexton PJ, Peterson CM, Boote KJ, White JW. Early-season growth in relation to region of domestication, seed size, and leaf traits in common bean. *Field Crops Research*. 1997; 52(1–2):69–78. [https://doi.org/10.1016/S0378-4290\(96\)03452-1](https://doi.org/10.1016/S0378-4290(96)03452-1)