

RESEARCH ARTICLE

Population structure and diversity of common bean (*Phaseolus vulgaris* L.) landraces in the Peruvian Amazon

Jorge Tobaru¹, Marvin Barrera-Lozano², Carlos D. Vecco-Giove², Jorge Peláez², Mack Pinchi³, Yesenia Vargas¹, Jorge Chuquillanqui¹, Julian Soto-Torres⁴, Ronald Robles⁴, David Saravia⁵, Cesar Petroli⁶, Jorge D. Etchevers⁷, Rodomiro Ortiz^{8*}, Raul Blas¹

1 Universidad Nacional Agraria La Molina (UNALM), Facultad de Agronomía, Programa de Leguminosas de Grano y Oleaginosas, La Molina-Lima, Perú, **2** Universidad Nacional de San Martín, Facultad de Ciencias Agrarias (UNSM), Tarapoto-San Martín, Perú, **3** Universidad Nacional de Ucayali (UNU), Facultad de Ciencias Agropecuarias, Pucallpa-Ucayali, Perú, **4** International Potato Center (CIP), La Molina-Lima, Peru, **5** Instituto Nacional de Innovación Agraria – INIA, La Molina-Lima, Peru, **6** International Maize and Wheat Improvement Center (CIMMYT), El Batán, Texcoco, Mexico, **7** Colegio de Postgraduados, Texcoco-Mexico, **8** Swedish University of Agricultural Sciences (SLU), Department of Plant Breeding, Alnarp, Sweden

* Rodomiro.ortiz@slu.se



OPEN ACCESS

Citation: Tobaru J, Barrera-Lozano M, Vecco-Giove CD, Peláez J, Pinchi M, Vargas Y, et al. (2026) Population structure and diversity of common bean (*Phaseolus vulgaris* L.) landraces in the Peruvian Amazon. PLoS One 21(7): e0332680. <https://doi.org/10.1371/journal.pone.0332680>

Editor: Santosh Gudi, North Dakota State University, UNITED STATES OF AMERICA

Received: September 2, 2025

Accepted: June 5, 2026

Published: July 20, 2026

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0332680>

Copyright: © 2026 Tobaru 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,

Abstract

Common bean (*Phaseolus vulgaris* L.) is a globally important grain legume, yet the genetic diversity of landraces from the Peruvian Amazon remains poorly characterized. This study aimed to assess the population structure and genetic diversity of Amazonian common bean by integrating morphological and genome-wide SNP data. A total of 476 accessions were phenotyped using 18 morphological traits, and 647 accessions were genotyped with 23,050 high-quality DArTseq SNP markers. Cluster and population structure analyses consistently identified two major genetic groups corresponding to the Andean ($n=284$) and Mesoamerican ($n=363$) gene pools, with further subdivision into four subgroups at $K=4$. Principal Coordinate Analysis supported this structure, with the first axis explaining 62.62% of the variation in the molecular data. Genetic diversity was high, with 94.9% of SNPs successfully mapped and more than 94% of loci polymorphic in both groups. However, observed heterozygosity was low ($H_o=0.035\text{--}0.041$) compared to expected heterozygosity ($H_e\approx 0.147\text{--}0.148$), and inbreeding coefficients were high ($F_{IS}=0.52\text{--}0.54$), consistent with the predominantly self-pollinating mating system. Hill number analyses indicated similar allelic richness between groups ($q=0$: $\alpha\approx 1.95\text{--}1.97$) and low differentiation in allele presence ($\beta=1.04$), whereas differentiation increased when allele frequencies were considered ($q=2$: $\beta\approx 1.4\text{--}1.5$). Morphological analyses revealed substantial phenotypic variation, including flowering time (32–75 days), pod length (6.5–17.3 cm), and number of locules per pod (2–11), and identified four distinct phenotypic clusters. No redundant accessions were detected. These results indicate that Amazonian

and reproduction in any medium, provided the original author and source are credited.

Data availability statement: Data is available at Figshare: Blas, Raul (2026). Morphological characterization data of common bean from Peruvian Amazon. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.32801379.v1> Blas, Raul (2026). DArTseq_SNP_647x23050 markers_common_bean. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.32801619.v1>.

Funding: This work was supported, in whole by Peruvian government funding-PROCIENCIA, E041 - Proyectos de Investigación Básica y Aplicada (PE501079155-2022), through the “Genómica del frijol *Phaseolus* spp. asociada con la diversidad cultural y ecosistémica en la Amazonia peruana” project.

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

landraces harbor high levels of genetic and phenotypic diversity structured by gene pool origin and local adaptation. The integration of morphological and SNP data provides a robust framework for germplasm characterization and highlights the importance of conserving these genetic resources for future breeding and sustainable agriculture.

Introduction

The common bean (*Phaseolus vulgaris* L.) an annual leguminous self-pollinated crop is the most widely cultivated species of the genus *Phaseolus* and a major source of dietary vegetal protein and income worldwide; which is a diploid ($2n = 2x = 22$) species and its genome size according to the first telomere-to-telomere (T2T) genome assembly was 560.30 Mb and predicted 29,925 protein-coding genes [1]. The origin of *P. vulgaris* remains controversial, separating it from Mesoamerica and the Andean region. However, recent research using molecular markers, including SNPs, shows that this species' origin is Mesoamerica [2–5]. Genetic and phylogeographic evidence indicate three distinct wild gene pools: 1) Mesoamerica (Mexico, Central America, Colombia, and Venezuela), 2) the Andes (southern Peru, Bolivia, and Argentina), and 3) northern Peru/Ecuador; nevertheless, the three common bean genetic groups belong to the primary gene pool [6]. From those, two main domesticated groups emerged, belonging to the Mesoamerican and Andean gene pools. This domestication process has been done independently in the Andes and Mesoamerica, which are associated with two Andean and Mesoamerican gene pools, respectively [7,8]. Both types of beans were introduced to other continents such as Europe, Africa, and Asia, where their widespread acceptance has led to debate about whether Europe can be a secondary center of diversification [9].

At the present, it is a staple crop in many parts of the world, most widely cultivated legume species worldwide, mainly due to its nutritional composition, being rich in amino acids, minerals, and vitamins—especially lysine, iron, zinc and folic acid [10]; presents medicinal attributes such as the prevention of cardiovascular diseases [11]; and has an ecological importance, it fixes atmospheric nitrogen reducing all the environmental costs involved in the acquisition of chemical fertilizers [12]. For this reason, the United Nations declared 2016 as “the international year of legumes” in its 68th session [13], the common bean being within the legumes, considered one of the essential crops to promote food security. Common bean has thus become an important source of income and nutrients for small farmers in the Caribbean, Latin America, Asia, and Africa, which account for 77% of global production [14–16]. In the case of Brazil, a predominantly Amazonian region, the Mesoamerican gene pool is found in a greater proportion (79%) and the Andean gene pool in a lesser proportion (21%) [4].

In Peru, the common bean is cultivated across diverse agroecological zones, ranging from coastal valleys to highland and Amazonian regions. According to national statistics [17], production covers approximately 42,575 hectares annually, with San Martín, Piura, and Cajamarca being the leading regions; with around 14% of this

area located in the Amazon region, 2.0% less than the area recorded in 2023. Beyond its economic relevance, the crop plays a crucial role in household nutrition and local food security. Despite its importance, most studies on common bean in Peru have focused on commercial cultivars grown in the highlands and coastal regions, emphasizing agronomic practices and genetic improvement using both local and introduced cultivars. Conversely, there is limited research conducted on the cultivars grown by farmers in indigenous communities of the Peruvian Amazon rainforest. According to the Genesys database, Peru holds 3,497 accessions of common beans, of which 2,838 (81%) are classified as traditional cultivars, 516 (15%) are results from breeding programs, and 143 (4%) correspond to wild or near-wild forms [18]. The International Center for Tropical Agriculture (CIAT) holds the global reference collection, while the Peruvian National Institute for Agrarian Innovation (INIA) maintains the national bean germplasm bank, which officially has 58 accessions of beans and 146 accessions of popping bean, mainly from the Andean region [19,20]. Despite these efforts, landraces and wild relatives from the Peruvian Amazon remain underrepresented, leaving a gap in the documentation of local agrobiodiversity.

Research on the diversity of cultivated common bean in Peru remains limited and has largely focused on specific bean types or restricted geographic areas. Most previous studies have emphasized popping beans (*ñuña*, in the Quechua language), employing phenotypic and molecular markers to characterize their diversity and differentiation [21,22], or have concentrated on isolated localities rather than addressing diversity at a broader regional scale [23]. In the central Peruvian Amazon, the cultivated and sociocultural diversity of common bean has been documented, highlighting the role of Indigenous communities and traditional knowledge in shaping bean diversity through local management practices and cultural preferences [24]. More recently, the agromorphological characterization of 58 *Phaseolus* spp. accessions collected in the Amazonas region identified several promising genotypes, with grain yields exceeding 2.77 t ha⁻¹, underscoring their potential value for use in breeding programs [25]. However, these studies have generally relied on limited sample sizes or phenotypic evaluations, constraining a comprehensive understanding of genome-wide diversity patterns. In contrast, high-throughput genotyping using DArTseq single nucleotide polymorphism (SNP) markers has been successfully applied to characterize genetic diversity and population structure in other legume crops, including cowpea [26], chickpea [27], and common bean germplasm from Ethiopia [28]. These studies demonstrate the power of DArTseq technology to resolve genetic relationships and identify valuable diversity for crop improvement, yet its application to common bean landraces from the Peruvian Amazon remains largely unexplored. The Peruvian Amazon harbors a wide array of common bean landraces and wild relatives, which are essential for preserving genetic diversity. The adaptation of common bean to the tropical rainforest conditions is particularly important due to the challenges associated with high humidity, variable soil fertility, and significant pathogen pressure. However, no systematic or representative collection has been established from the Peruvian Amazon, nor has a comprehensive analysis of the genetic diversity across its rainforest regions been conducted. As a result, little is known about the genetic diversity and origins of these cultivars, which may represent an untapped reservoir of alleles for crop improvement. This study aimed to assess the genetic diversity of common bean landraces collected in the Peruvian Amazon using both morphological descriptors and genome-wide SNP markers (DArTseq). By combining phenotypic and molecular approaches, we sought to clarify the structure of local germplasm and to generate information that supports both conservation strategies and future breeding efforts.

Materials and Methods

Sample Collection

A total of 1,025 common bean (*P. vulgaris*) accessions were recently collected from 10 rainforest regions of Peru: Amazonas, Cajamarca, San Martín, Huánuco, La Libertad, Loreto, Ucayali, Cerro de Pasco, Junín, and Cusco across the Peruvian Amazon, including Indigenous farming systems and traditional local markets (Fig 1). From this collection, a subset of 633 accessions—representative from Amazonian region—was selected for morphological and molecular characterization. Also, it was added 34 accessions for genotyping as a reference genetic material belonging to both Mesoamerican and Andean gene pools, looking for the relationships between accessions and to help to define gene pool group; from them

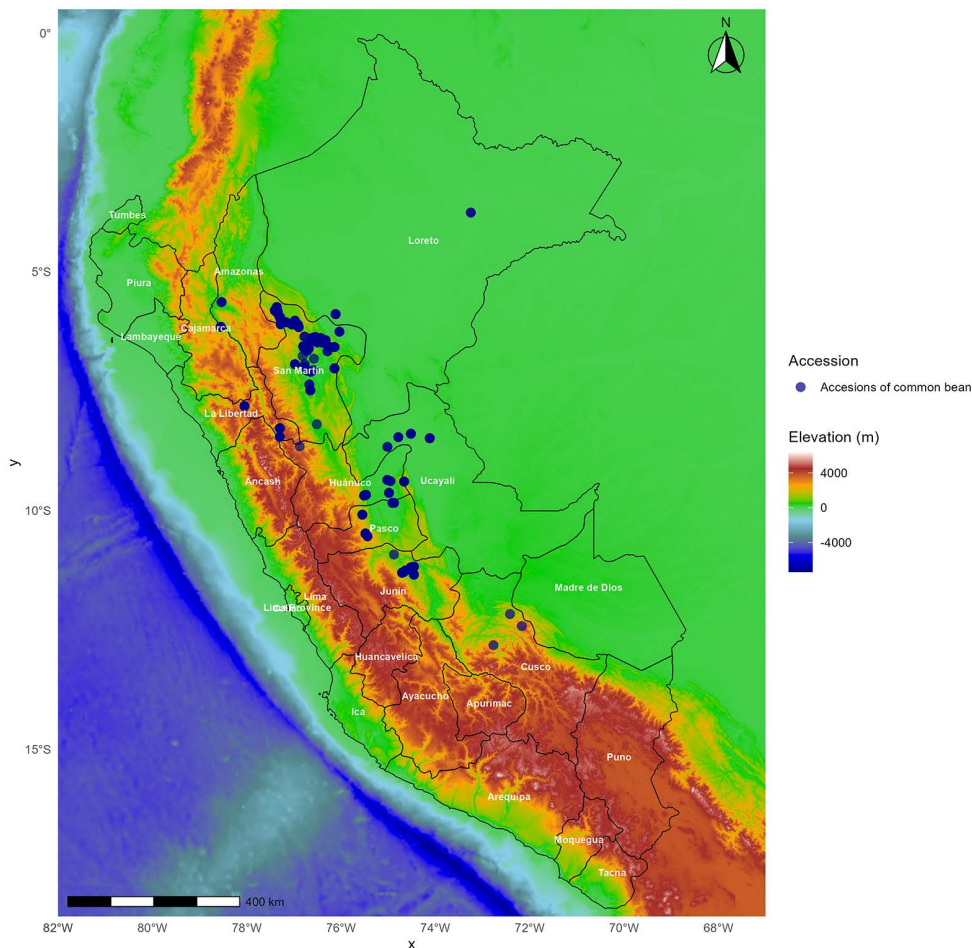


Fig 1. Distribution of Common bean accession from Peruvian Amazonia.

<https://doi.org/10.1371/journal.pone.0332680.g001>

26 are from Peru (12 from Coast, 14 from Highlands), 7 from CIAT (2 Andean y 5 Mesoamerican gene pool), and 1 from Mexico which was considered as a reference for Mesoamerican gene pool.

The map was elaborated using R program, according the following methodological approach for creating the map of bean accessions in Software R consisted of the following sequence: 1) obtaining elevation data (DEM) for the Peruvian territory using the `get_elev_raster()` function from the `elevatr` package; 2) loading political boundary vector layers via `rnatuarearth`; 3) importing and converting collected bean occurrence points—stored in clipboard data—into a spatial format (`sf`); and 4) integrating all layers into a two dimensional map with `ggplot2`, overlaying the shaded relief (elevation raster with a customized color palette), departmental borders, collection locations and cartographic elements such as a north arrow and scale bar. Thereafter, the map was exported as a TIF image.

Morphological Characterization

The experimentation was conducted in the 2024 cropping season at the experimental field of Universidad Nacional Agraria La Molina (UNALM), Lima, Peru. The density of sowing was in a 2.5 m row: 35 cm between plants and 50 cm between rows, with up to three seeds per hole (in total by experimental unit, 7 holes and 21 plants). Phenotypic data were recorded for each accession following the standardized common bean (*Phaseolus vulgaris* L.) descriptor guidelines

established by the International Board for Plant Genetic Resources [29]. Traits evaluated included plant growth habit and pod characteristics (shape, color, size, and presence of sutures or fiber). Throughout the growing cycle, a total of 18 traits were recorded—3 quantitative and 15 qualitative—covering vegetative, floral, pod, and seed characteristics (Table 1); such as seed shape, size, color, and plant growth habit were evaluated according to standard descriptors. For each accession, five plants were evaluated, and from each plant, 10 pods and ten seeds were analyzed. Pods were collected at full physiological maturity. The evaluated descriptors and their corresponding measurement criteria were as follows:

Growth habit: Determined based on plant architectural characteristics, using the descriptive categories and reference diagrams provided in the descriptor manual.

Table 1. Quantitative and qualitative traits recorded in the *Phaseolus vulgaris* accessions.

Order	Descriptor title	Descriptor data type	Coding ¹
1	Plant type/ Growth habit	Coded	1 Determinate bush; 2 Indeterminate bush; 3 Indeterminate semi-climber or prostrate (with many lateral guides); 4 Indeterminate climber
2	Days to flowering	Numeric	Number of days from emergence to stage where 50% of plants have set flowers
3	Color of standard	Coded	In freshly opened flowers: 1 White; 2 Green; 3 Lilac; 4 White with lilac edge; 5 White with red stripes; 6 Dark lilac with purple outer; 7 Dark lilac with purplish spots; 8 Carmine red; 9 Purple; 10 Light lilac; 11 Dark lilac
4	Color of wings	Coded	In freshly opened flowers: 1 White; 2 Green; 3 Lilac; 4 White with carmine stripes; 5 Strongly veined in red to dark lilac; 6 Plain red to dark lilac; 7 Lilac with dark lilac veins; 8 Purple; 9 Light lilac; 10 Dark lilac
5	Pod color	Coded	From fully expanded immature pod: 1 Dark purple; 2 Carmine red; 3 Purple stripe on green; 4 Carmine stripe on green; 5 Pale red stripe on green; 6 Dark pink; 7 Normal green; 8 Shiny green; 9 Dull green to silver grey; 10 Golden or deep yellow; 11 Pale yellow to white.
6	Pod length	Numeric	Average length in centimeters of the largest fully expanded immature pods from five random normal plants
7	Pod cross-section	Coded	From fully expanded immature pod: 1 Very flat; 2 Pear shaped; 3 Round elliptic; 4 Figure of eight
8	Pod curvature	Scale	Fully expanded immature pod: 3 Straight; 5 Slightly curved; 7 Curved; 9 Recurving
9	Locules per pod	Numeric	Number of locules from longest pod of five random normal plants
10	Pod wall fibers	Scale	3 Strongly contracting (at dry maturity adhering around seed)5 Leathery podded (dry pods will not spontaneously open)7 Excessive shattering (with strong twisting of dry pods)
11	Pod apex position	Coded	1 = marginal; 2 = non-marginal; 3 = other.
12	Pod apex orientation	Coded	1 = upwards (dorsal direction); 2 = straight; 3 = downwards (ventral direction).
13	Pod color at the physiological maturity	Coded	1 Dark purple; 9 Silver; 10 Golden or deep yellow; 11 Light yellow to white; 12 Deep yellow with purple mottling or strips; 13 Pale yellow with purple mottling or stripes; 14 Pale yellow with red mottling or stripes
14	Seed coat pattern	Coded	1 Constant mottled; 2 Striped; 3 Rhomboid spotted; 4 Speckled; 5 Circular mottling; 6 Marginal color pattern; 7 Broad striped; 8 Bicolor; 9 Spotted bicolor; 10 Absent or whole color
15	Seed coat lighter color	Text	1 Black; 2 Brown, pale to dark; 3 Maroon; 4 Grey, brownish to greenish; 5 Yellow to greenish; 6 Pale-cream to buff; 7 Pure white; 8 Whitish; 12 Red; 13 Pink; 14 Purple; 15 Carmine red or wine; 16 Dark yellow
16	Seed coat darker color	Coded	1 Black; 2 Brown, pale to dark; 3 Maroon; 4 Grey, brownish to greenish; 5 Yellow to greenish; 6 Pale-cream to buff; 7 Pure white; 8 Whitish; 12 Red; 13 Pink; 14 Purple; 15 Carmine red or wine; 16 Dark yellow
17	Seed brightness	Scale	3 Matt; 5 Medium; 7 Shiny
18	Seed shape	Coded	Taken from middle of pod:1 Round; 2 Oval; 3 Cuboid; 4 Kidney shaped; 5 Truncate fastigiate

¹Adapted from IBPGR—International Board for Plant Genetic Resources

Days to flowering: Recorded as the number of days from sowing to the opening of the first floral bud in at least 50% of the evaluated plants per accession.

Standard color: Assessed on freshly opened flowers collected from the inflorescence at the fourth node of each evaluated plant, using the standardized descriptor color chart.

Wing color: Recorded from freshly opened flowers collected from the inflorescence at the fourth node of each evaluated plant, following the descriptor color chart.

Green pod color: Evaluated on 10 fully developed immature pods per accession, using the descriptor color chart.

Pod length: Calculated as the mean pod length (cm) based on measurements of 10 pods collected from five representative plants per accession.

Pod cross-section: Determined by examining a transverse section of 10 fully developed immature pods at the level of the second seed from the distal end, following the descriptor reference chart.

Pod curvature: Assessed on 10 fully developed immature pods per accession, according to the descriptor description chart.

Number of locules: Calculated as the mean number of locules determined from the longest pod of each of the 10 pods collected from five plants per accession.

Pod wall fiber: Evaluated on 10 pods collected from five plants per accession, using the descriptor model to classify fiber development.

Pod apex position: Recorded from 10 pods collected from five plants per accession, following the descriptor model.

Pod apex orientation: Assessed on the same set of 10 pods per accession, based on the descriptor model.

Pod color at physiological maturity: Evaluated on 10 dry pods at harvest, using the standardized descriptor color chart.

Seed coat pattern: Recorded from 10 seeds per accession, each taken from the seed closest to the pod apex, using the descriptor pattern chart.

Seed coat lighter color: The primary (lighter) seed coat color was recorded from 10 seeds per accession, each taken from the seed closest to the pod apex, using the descriptor color chart.

Seed coat darker color: The secondary (darker or background) seed coat color was recorded from the same 10 seeds per accession, following the descriptor color chart.

Seed brightness: Assessed as seed coat luster on 10 seeds per accession, each taken from the seed closest to the pod apex.

Seed shape: Recorded from 10 seeds taken from the middle portion of each pod, following the descriptor classification scheme.

Morphological data analysis

Before statistical analyses, quantitative variables were summarized by calculating arithmetic means, while qualitative variables were summarized using their modal values. The resulting data matrix was subsequently used for both univariate and multivariate analyses. Quantitative traits were described using boxplots and summary statistics, including mean, minimum, and maximum values, to characterize the range and distribution of phenotypic variation. Qualitative traits were coded using standardized descriptor scales and graphically represented as frequency histograms to visualize their distribution across accessions. A principal component analysis (PCA) was used to examine the association between the analyzed traits and the similarity among accessions, and Multiple Correspondence Analysis (MCA) was performed to associate qualitative variables with the groups. PCA was performed with all quantitative traits and qualitative categorized values. Additionally, for all traits, a Gower's distance was computed for clustering analysis by using the Ward. D2 method. All statistical analyses and graphical visualizations were performed using R version 4.5.1 [30] and RStudio version 2025.05.1 + 513 [31], a programming language for statistical computing and data visualization, using R Base and packages as "FactoMineR" [32], "factoextra" [33], "ggplot2" [34], and "randomForest" [35].

Genotyping

DNA isolation and sequencing. Seeds of each common bean accession were germinated in a shade house. Leaf samples were collected from two-week-old seedlings, immediately frozen at -80°C , and used for DNA extraction. Genomic DNA was extracted from frozen leaf tissue using CTAB protocol. DNA quality and concentration were initially assessed using a NanoDrop spectrophotometer and subsequently quantified with a Qubit fluorometer (Life Technologies, Carlsbad, CA, USA). DNA Integrity was verified via 0.8% agarose gel electrophoresis. The DNA concentration was adjusted to $50\text{--}100\text{ ng }\mu\text{l}^{-1}$, and samples were submitted for genotyping using the high-throughput DArTSeq™ genotyping platform at SAGA-CIMMYT(Mexico), following the protocols described by Sansaloni et al. [36] and Kilian et al. [37]. This technology uses a complexity reduction approach based on a digestion/ligation reaction, employing a combination of two restriction enzymes, PstI and Msi, to enrich genomic representations with single-copy sequences. PstI-compatible adapters were ligated for subsequent PCR amplification, along with sample-specific barcodes added to the digested fragments. Equimolar amounts of amplified products from each sample were pooled and sequenced on an Illumina Novaseq 6000 platform. Sequence data were processed using proprietary DArT analytical pipelines to filter low-quality reads and assign sequences to individual samples based on barcode information.

SNP calling and genotypic data filtering

Molecular markers discovery was performed using an analytical pipeline (DARTsoft14) developed by DArT company. This software identifies two types of markers: SilicoDART (presence/absence) and SNP markers [38]. Marker scoring and reproducibility were assessed using technical replicates, and only markers with high call rates and reproducibility were retained. Both markers were aligned to the *Phaseolus vulgaris* YP4 reference genome [1] to identify the chromosome positions. Initially, we received 80,079 DArTSeq-derived SNP markers from SAGA-CIMMYT, which were polymorphic across common bean genotypes analyzed. Markers with unknown position were first removed from the analysis. Marker parameters such as call rate and minimum allele frequency (MAF) were calculated using dartR package in R version 4.5.1 [39]. Accordingly, SNP markers with $\geq 50\%$ call rate and MAF of $\geq 0.5\%$ were retained for further analysis. In this study, genotypes with missing data above $\geq 30\%$ were removed from the analysis. For SilicoDART markers (presence/absence), the initial report of 462,250 markers was filtered to retain markers with a call rate $\geq 90\%$ and MAF of $\geq 0.5\%$. Markers with unknown position were also removed from the final analysis.

Genetic relationships and diversity metrics

Pairwise genetic distances among accessions were calculated using the scaled Euclidean distance method, as implemented in the gl.dist.ind function of the dartR package. Genetic diversity indices, including expected heterozygosity (H_e), observed heterozygosity, polymorphic information content (PIC), major allele frequency, and inbreeding coefficient (F_{is}), were computed using functions also available in dartR package.

To infer relationships among accessions, we constructed a neighbor-joining tree, based on the previous Euclidean distance matrix, in DARwin v6.2.1. [40]. Principal Coordinate Analysis (PCoA) was conducted using the gl.pcoa function in dartR to visualize genetic relationships among accessions in reduced dimensions.

Population structure and genetic diversity analysis

Population structure was inferred using fastSTRUCTURE, a framework for inferring population structure from a large SNP genotype dataset [41]. Analyses were conducted for a range of K values (number of assumed populations) from 1 to 10. The ancestry proportions for each individual (Q-matrix) were visualized alongside the neighbor-joining tree and combined with the geographical passport data to contextualize cluster assignments.

Results

Nominal variability of common bean

Based on passport data collected during fieldwork, a total of 74 different common names were recorded for common bean from Peruvian Amazonia. The ten most frequently cited cultivars were Panamito, Waska, Ashpa, Pinto, Canario, Poroto, Ñuña (popping bean), Shingo, Manteca, and Awisho ([Table 2](#)). Of these, Panamito and Canario are commercial cultivars likely introduced from the coastal region. These are predominantly cultivated in higher-altitude zones by migrant farmers from the coast and highlands. The remaining cultivars are traditional landraces distributed across both the lowland and upland Amazon. The most common cultivar collected was Panamito, representing 15.4% of all accessions. This was followed by local landraces such as Waska (12.4%), Ashpa (11.8%), and Pinto (6.5%). The four bean cultivars most widely cultivated in the Peruvian Amazon collectively accounted for only 5.4% of the total cultivated material. In contrast, 17 cultivars (representing 23% of the total) exhibited low frequencies (1–4%), whereas the majority—70 cultivars—were cultivated at frequencies below 1%, together comprising 71.6% of the recorded landraces. This highly skewed frequency distribution indicates that most Amazonian bean landraces are maintained under localized and low-intensity cultivation systems. Such a pattern underscores their vulnerability to genetic erosion and highlights the need to explicitly incorporate rarity and spatial restriction into strategies for the conservation, management, and effective utilization of Amazonian common bean germplasm.

Morphological variability

We describe the characterization of 476 accessions (from 667 sown); the remaining 191 accessions did not progress in the field trial. Based on qualitative and quantitative morphological descriptors, we estimated the variability of the collected accessions. Substantial phenotypic variability was observed among the accessions, particularly in traits such as seed color and growth habit, reflecting both natural selection and farmer-driven selection ([Fig 2](#)).

The boxplot for quantitative traits shows high variability. Flowering time ranged from 32 to 75 days, showing that some common bean landraces have short vegetative growth periods, which would be important for use in the development of early-maturing cultivars. Pod length ranged from 6.5 to 17.3 cm, while locules per pod ranged from 2 to 11, which are very important sources for common bean plant breeders because these characteristics offer important yield components to develop improved cultivars for the tropical rainforest ([Fig 3](#)).

The histogram analysis of qualitative traits revealed substantial phenotypic variation in flower, pod, and seed characteristics, including color, size, and shape ([Fig 4](#)). All growth habit categories defined in the descriptor list were represented. Indeterminate bush types were the most frequent, followed by determinate bush types, whereas indeterminate prostrate and indeterminate climbing forms occurred at low frequencies. This distribution suggests a predominance of plant architectures commonly associated with shorter vegetative growth cycles in environments such as the Peruvian Amazon.

Flower color traits showed clear patterns of variation. White was the predominant standard color, followed by green, lilac, white with lilac margins, and purple. Wing color exhibited a similar distribution, with white being the most common, followed by green, lilac, and purple.

Pod color was mainly green, with additional variation including purple-striped green, carmine-striped green, and pale red-striped green pods. Pod cross-section was predominantly pear-shaped, followed by round-elliptic, very flat, and Fig-eight forms. Slightly curved pods were most frequent, followed by straight and curved pods.

Regarding pod wall fibers, excessive shattering—characterized by strong twisting of dry pods—was the most common condition, followed by leathery pods, in which dry pods do not open spontaneously. The pod apex position was mainly marginal, and apex orientation was predominantly downward (ventral).

At physiological maturity, pod color was primarily light yellow to white, followed by golden or deep yellow, pale yellow with purple mottling or stripes, deep yellow with purple mottling, and pale yellow with red mottling.

Table 2. Nominal variability of common bean from Peruvian Amazonia.

Common name	Number	Frequency	Percentage
Panamito	143	0.15409483	15.41
Waska	115	0.12392241	12.39
Ashpa	109	0.11745690	11.75
Pinto	60	0.06465517	6.47
Canario	36	0.03879310	3.88
Poroto	36	0.03879310	3.88
Ñuña	34	0.03663793	3.66
Shingo	31	0.03340517	3.34
Manteca	27	0.02909483	2.91
Awisho	23	0.02478448	2.48
Ojo de cabra	22	0.02370690	2.37
Cambio-90	20	0.02155172	2.16
Chaucha	20	0.02155172	2.16
Norteño	20	0.02155172	2.16
Cápsula	14	0.01508621	1.51
Vaca paleta	14	0.01508621	1.51
Chiclayo	11	0.01185345	1.19
Pozuzo	11	0.01185345	1.19
Ashpilla	10	0.01077586	1.08
Garrapata	10	0.01077586	1.08
Mantequilla	10	0.01077586	1.08
Bayo	9	0.00969828	0.97
Caballero	9	0.00969828	0.97
Ninaporoto	9	0.00969828	0.97
San Pablino	9	0.00969828	0.97
Ucayalino	9	0.00969828	0.97
60 días	8	0.00862069	0.86
Chuncho	7	0.00754310	0.75
Pintado	6	0.00646552	0.65
Vaquita	6	0.00646552	0.65
Leche	5	0.00538793	0.54
Frijol local	5	0.00538793	0.54
Cuarentón	4	0.00431034	0.43
Dark	4	0.00431034	0.43
Reclina	4	0.00431034	0.43
Blanco	3	0.00323276	0.32
Chaoca	3	0.00323276	0.32
Nube	3	0.00323276	0.32
Panamo	3	0.00323276	0.32
Burra	2	0.00215517	0.22
Caraota	2	0.00215517	0.22
Chinito	2	0.00215517	0.22
Cumbeño	2	0.00215517	0.22
Frejol de suelo	2	0.00215517	0.22
Pavita	2	0.00215517	0.22
Región	2	0.00215517	0.22

(Continued)

Table 2. (Continued)

Common name	Number	Frequency	Percentage
Sangre de toro	2	0.00215517	0.22
Amarillo	1	0.00107759	0.11
Arverja	1	0.00107759	0.11
Bayito	1	0.00107759	0.11
Bola	1	0.00107759	0.11
Camote	1	0.00107759	0.11
Canarillo	1	0.00107759	0.11
Comandante	1	0.00107759	0.11
Grande	1	0.00107759	0.11
Huallaguino	1	0.00107759	0.11
Huanuqueño	1	0.00107759	0.11
Huevo de pajarito	1	0.00107759	0.11
Lanche	1	0.00107759	0.11
Lentreja	1	0.00107759	0.11
Limón	1	0.00107759	0.11
Mani	1	0.00107759	0.11
Mocho	1	0.00107759	0.11
Nina panamito	1	0.00107759	0.11
Riñón	1	0.00107759	0.11
Rojo	1	0.00107759	0.11
Rundo	1	0.00107759	0.11
Seda	1	0.00107759	0.11
Tabaquero	1	0.00107759	0.11
Trujillano	1	0.00107759	0.11
Tumba-maíz	1	0.00107759	0.11
Vainita	1	0.00107759	0.11
Wayruro	1	0.00107759	0.11
Total	928	1.00000000	100.00

<https://doi.org/10.1371/journal.pone.0332680.t002>

Considerable diversity was observed in seed traits. All seed coat pattern states were detected, with whole color being the most frequent, followed by rhomboid-spotted, striped, speckled, broad-striped, and bicolor patterns. Variation was also evident in seed coat lighter color, predominantly yellow to greenish, followed by black, carmine red or wine, pure white, pale cream to buff, and gray. Seed coat darker color was mainly carmine red or wine and black, with additional classes including yellow to greenish, brownish to greenish, pale cream to buff, purple, and pink.

Seed brightness was predominantly intermediate, followed by matte and shiny types. Seed shape was mainly cuboid, followed by truncate-fastigiate, oval, round, and kidney-shaped forms.

Overall, the high level of qualitative trait diversity observed in this collection highlights its value as a genetic resource for the conservation, utilization, and improvement of common bean adapted to Amazonian environments.

Relationships and variability

Cluster analysis revealed the presence of four well-defined groups among common bean (*Phaseolus vulgaris* L.) accessions from the Peruvian Amazon (Fig 5). Group sizes were uneven, comprising 78, 97, 250, and 51 accessions for clusters 1–4, respectively. This clear partitioning indicates a structured pattern of morphological diversity

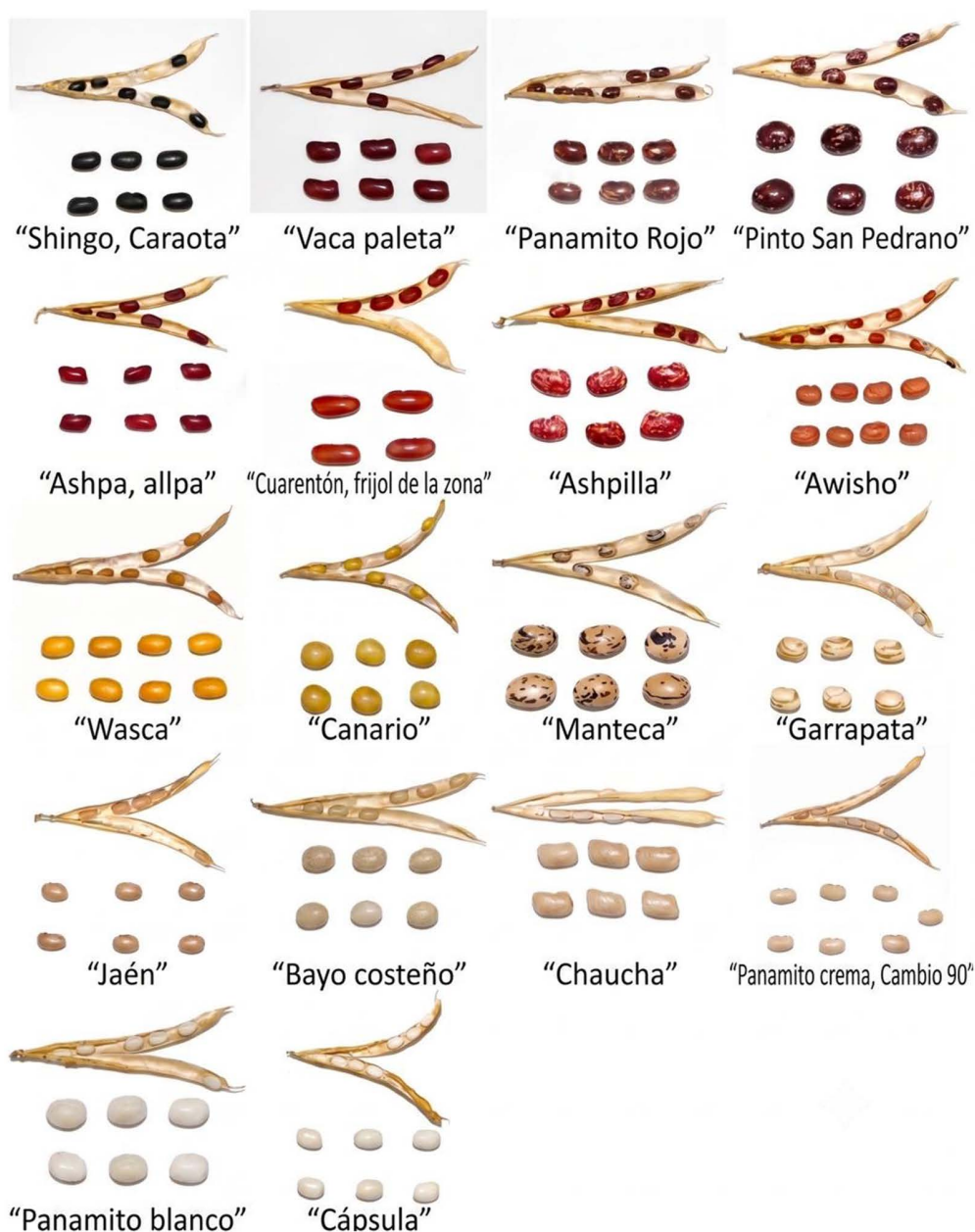


Fig 2. Morphological variability of common bean from Peruvian Amazonia, showing different forms, sizes, colors of the seed, and ripe pods.

<https://doi.org/10.1371/journal.pone.0332680.g002>

within these landraces. Cluster formation was driven by 18 discriminant descriptors—three quantitative and 15 qualitative—that captured substantial variation in plant architecture, phenology, and seed-related traits. The observed diversity reflects not only underlying genetic differentiation but also the influence of farmer-mediated selection and adaptation to contrasting agroecological conditions. The main characteristics defining each group are summarized below.

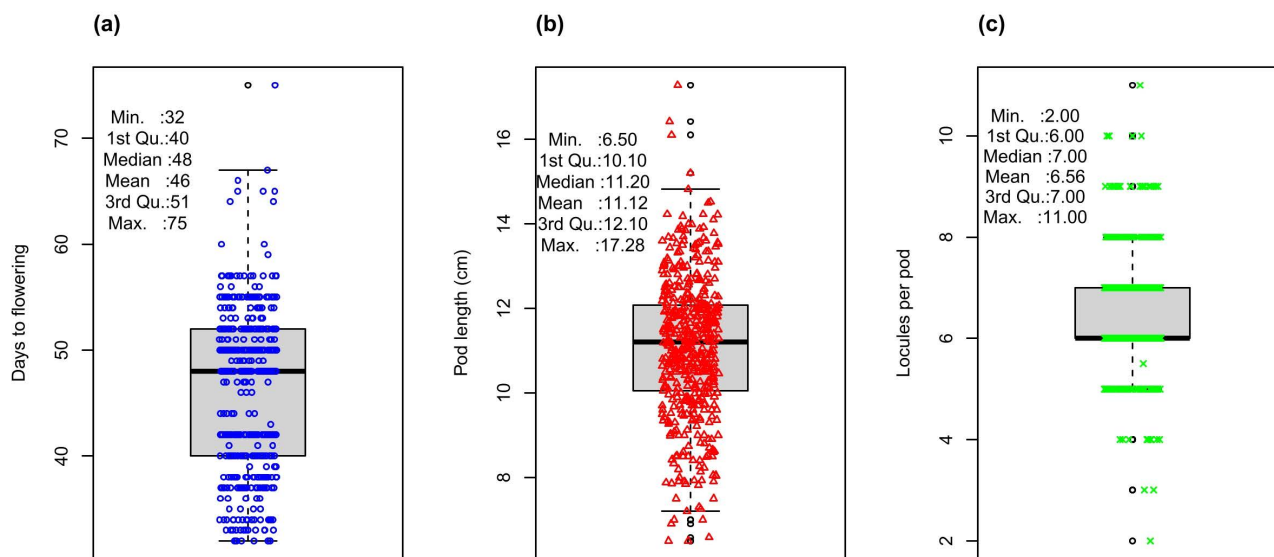


Fig 3. Boxplot of quantitative descriptors on 476 common bean accessions. a=Days to flowering, b=Pod length, and c=Number of locules per pod.

<https://doi.org/10.1371/journal.pone.0332680.g003>

Group 1: Short vegetative cycle

Early flowering and rapid physiological maturity characterize accessions in this group. Plants predominantly exhibit a determinate growth habit, reduced vegetative development, and relatively small leaflets, traits typically associated with early-cycle and low-input production systems.

Group 2: Climbing habit

This group was predominantly composed of accessions exhibiting an indeterminate climbing growth habit (Type IV), characterized by vigorous vine development, elongated internodes, and a strong climbing ability. These accessions tend to exhibit later maturity, extended vegetative growth, and a higher number of pods per plant, reflecting their suitability for longer growing cycles and supported cropping systems.

Group 3: Vegetative vigor

The accessions exhibited vigorous vegetative growth, characterized by dense canopy development, broad leaflets, and robust stems with high mechanical strength. Growth habits are mainly Type II or III, suggesting a favorable balance between vegetative biomass and reproductive allocation.

Group 4: Intermediate Morphological Types

This group includes accessions with intermediate values for quantitative descriptors. Differentiation within this cluster is primarily driven by qualitative traits, such as flower color and seed-coat patterning. These accessions appear to represent transitional or recombinant morphotypes, potentially acting as a genetic bridge between more extreme phenotypic classes.

The identified clusters were not strictly associated with geographic origin; however, certain accessions, such as Awisho and Ashpa, consistently grouped and were predominantly cultivated in lowland environments. This pattern suggests incipient morphological differentiation linked to ecological factors, particularly altitude-related environmental gradients. Overall, the structured diversity observed among these landraces highlights their considerable potential for use in breeding programs aimed at improving yield and adaptation, while simultaneously conserving the rich agrobiodiversity of the Peruvian Amazon.

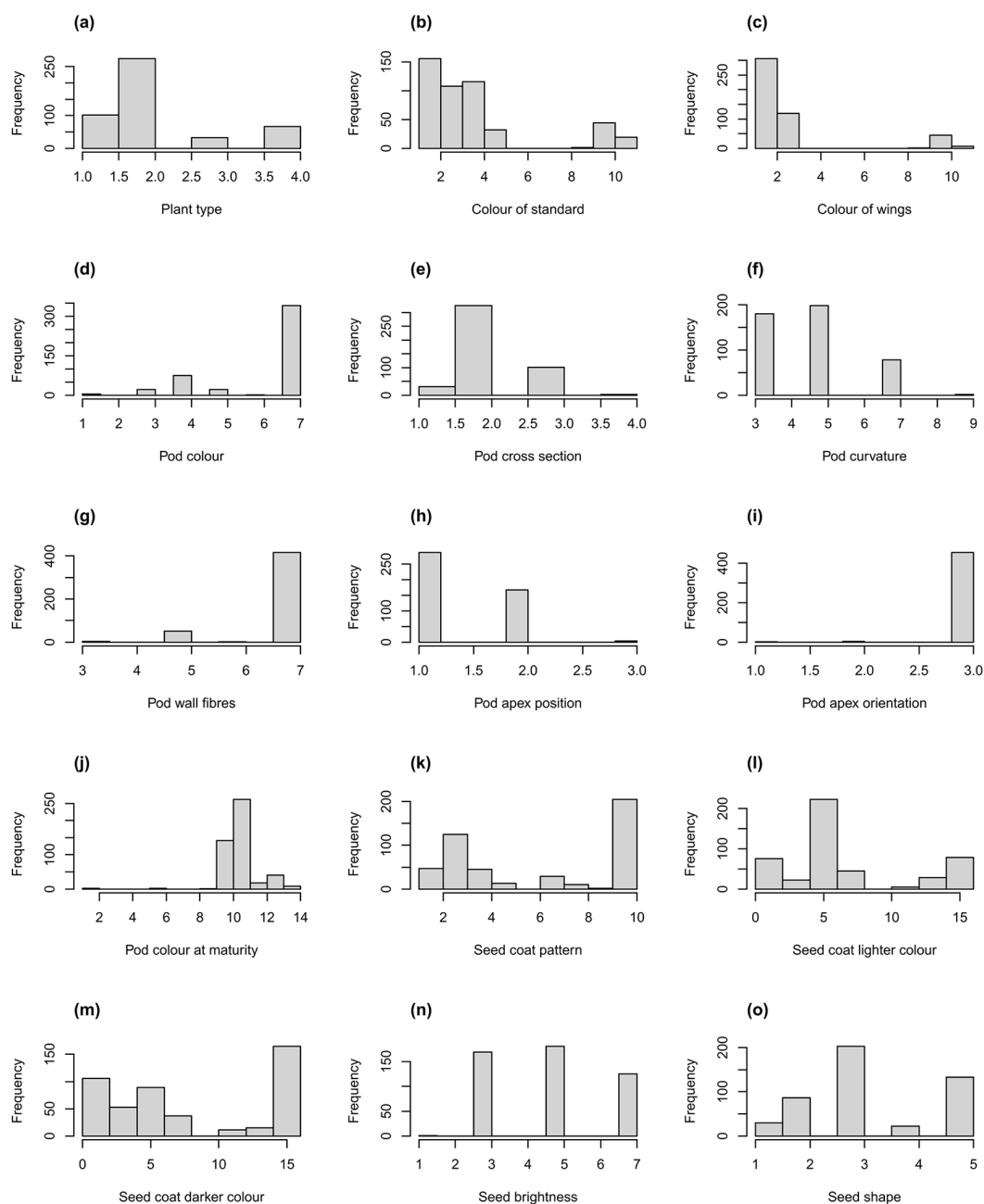


Fig 4. Histogram of 15 qualitative descriptors on 476 common bean accessions.

<https://doi.org/10.1371/journal.pone.0332680.g004>

The first two principal components (PC1 and PC2) explained 16.9% and 11.2% of the total phenotypic variation, respectively, accounting for a cumulative 28.1% (Fig 6). The relatively low proportion of variance captured by the first two components indicates a high level of morphological complexity among the evaluated accessions, with the remaining 71.9% of variation distributed across subsequent components. This pattern suggests that multiple traits with moderate contributions collectively shape phenotypic differentiation, rather than a few dominant variables.

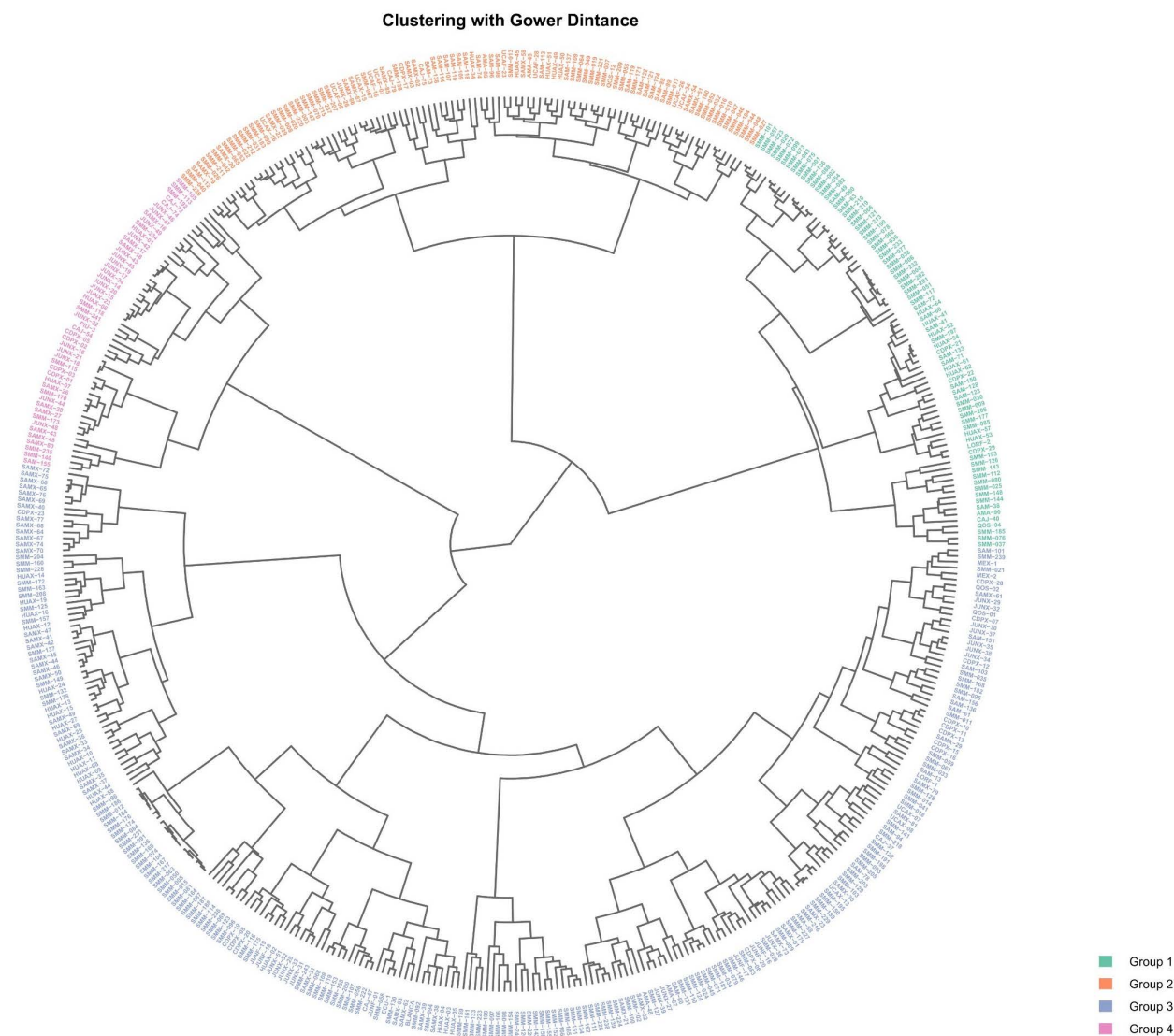


Fig 5. Tree of 476 common bean accessions according to 18 morphological descriptors. The color indicates the groups: group 1 = green, group 2 = orange, group 3 = grey, and group 4 = pink.

<https://doi.org/10.1371/journal.pone.0332680.g005>

The PCA biplot revealed the presence of four groups, broadly consistent with the clusters identified through hierarchical cluster analysis. However, clear separation based on morphological traits was evident only for groups 1 and 4. Group 1 was characterized by late flowering (approximately 50 days after sowing), excessive shattering of pod wall fibers, kidney-shaped and truncate-fastigiate seed forms, and a relatively high number of locules per pod (approximately 8). In contrast, group 4 comprised early flowering accessions (approximately 38 days), strongly contracting pod wall fibers, predominantly round to oval seed shapes, and a lower number of locules per pod (approximately 5.6).

In contrast, groups 2 and 3 showed substantial overlap in the multivariate space, with no clear morphological boundaries between them. This lack of separation suggests phenotypic admixture or shared trait combinations among these groups, potentially reflecting gene flow, parallel selection by farmers, or convergent adaptation to similar agroecological

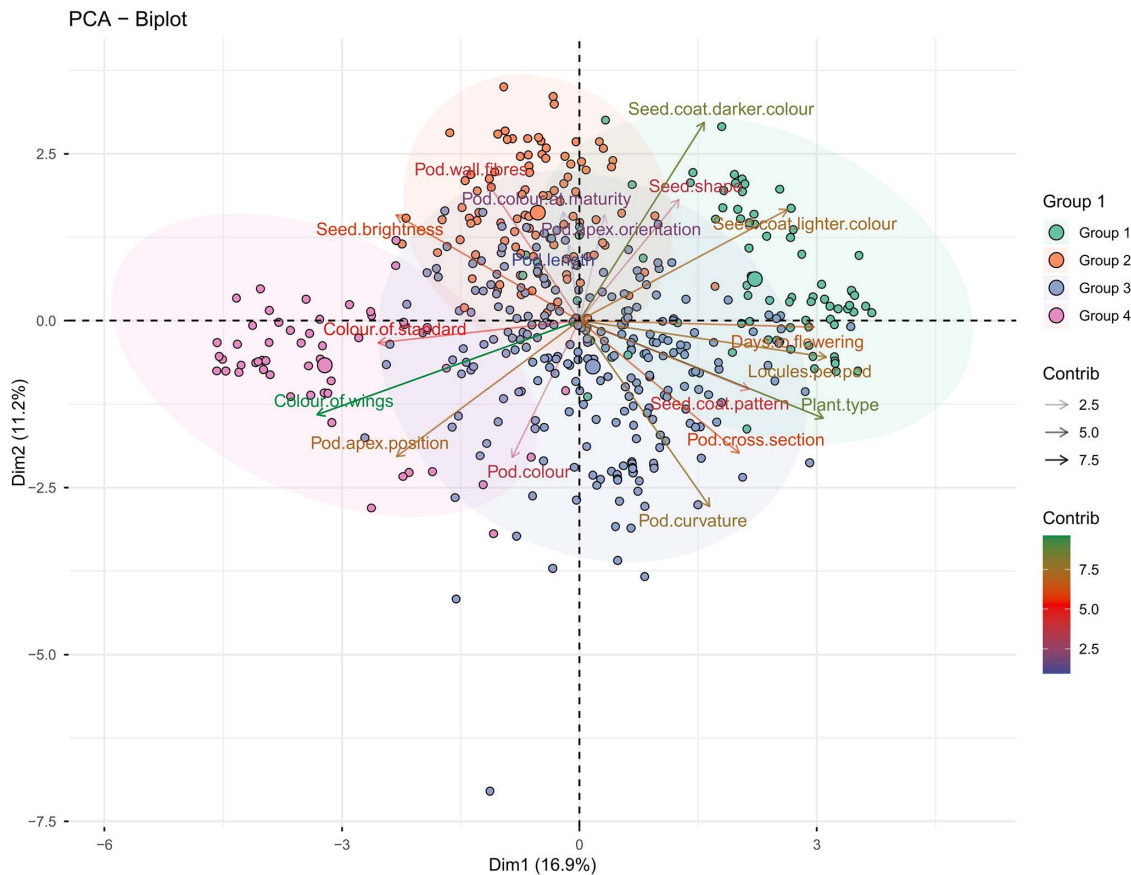


Fig 6. The first two PC biplots of 476 common bean accessions according to 18 morphological descriptors. The color indicates the groups: group 1 = green, group 2 = orange, group 3 = grey, and group 4 = pink.

<https://doi.org/10.1371/journal.pone.0332680.g006>

conditions. Overall, the PCA results support the cluster analysis while highlighting that only a subset of morphological traits contributes to clear phenotypic differentiation among the identified groups.

Multiple Correspondence Analysis (MCA), performed using the SNP-defined subgroups, showed that, in general, the subgroups were not distinguished by exclusive qualitative descriptors (Fig 7). This result indicates a high degree of overlap in qualitative morphological traits among genetically defined groups. An exception was observed for subgroup 1.2, which was associated with lilac flower coloration (standard and wings), light red pod color at physiological maturity, and a carmine seed coat color. This subgroup corresponds to the Andean gene pool, suggesting that, in this case, certain qualitative descriptors are partially congruent with the genetic structure revealed by SNP data.

SNP marker variation and distribution

From the initial set of 665 common bean accessions, 647 presented less than 30% missing data and were retained for downstream analyses. SNP markers were filtered to retain only those with a call rate of at least 50% and a minor allele frequency (MAF) of 0.5% or higher, yielding a filtered dataset of 24,283 high-quality SNPs (Fig 8, Table 3). Alignment to the common bean YP4 reference genome revealed that 23,050 SNPs (94.9%) were successfully mapped to chromosome positions.



Fig 7. Multiple Correspondence Analysis (MCA) of 476 accessions with 15 qualitative descriptors. The color indicates the groups and sub-groups according to the SNP-defined subgroups.

<https://doi.org/10.1371/journal.pone.0332680.g007>

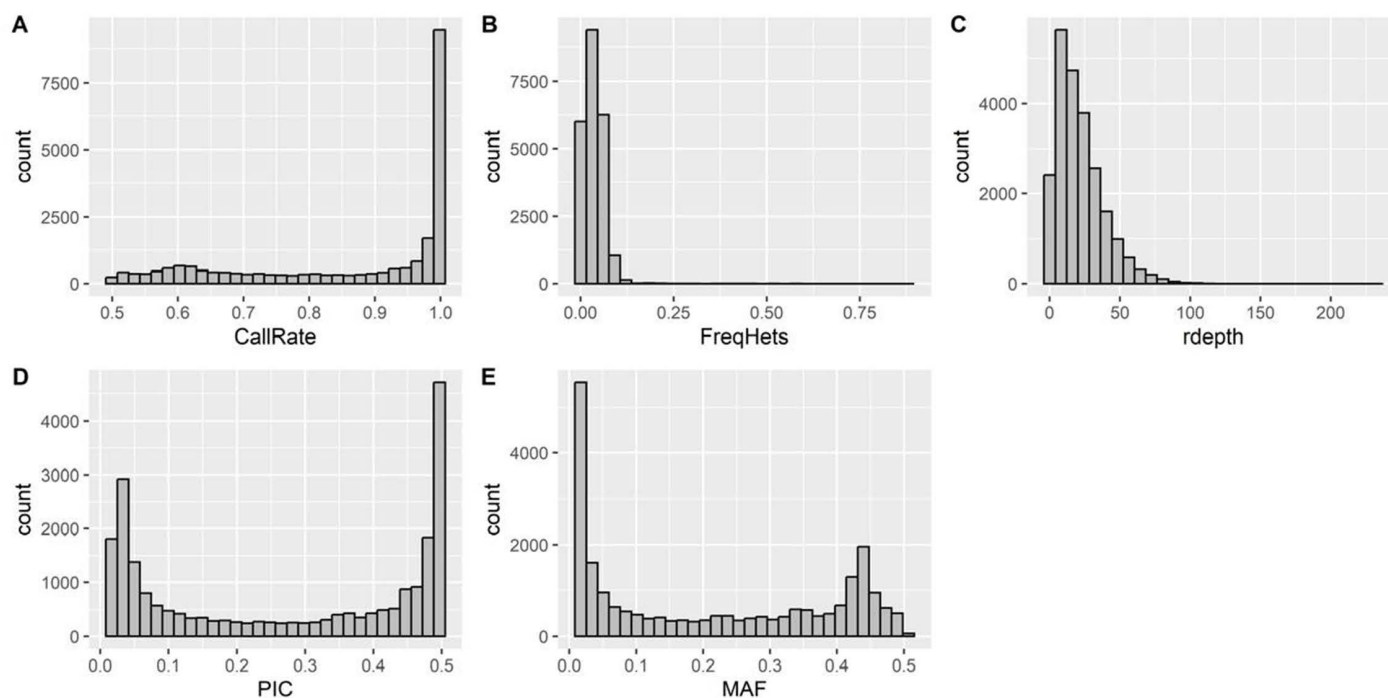


Fig 8. Histogram distribution of single nucleotide polymorphism (SNP) marker parameters. A) Call rate, B) Frequency of heterozygous genotypes, C) Read depth, D) Polymorphic Index Content (PIC), and E) Minimum allele frequency (MAF) on the filtered set of 23050 SNP markers.

<https://doi.org/10.1371/journal.pone.0332680.g008>

Table 3. Single nucleotide polymorphism (SNP) call rate report pre and post filtering.

	Reporting call rate by		Call rate post filtering	
	Individual	Locus	Individual	Locus
No. of loci	80079	80079	23050	23050
No. of individuals	665	665	647	647
Minimum	0.3658887	0.3443610	0.5007730	0.7078525
1st quartile	0.8411069	0.7015040	0.7140650	0.8309761
Median	0.8577155	0.9669170	0.9675430	0.8640781
Mean	0.8507188	0.8507188	0.8623075	0.8623075
3r quartile	0.8785449	0.9894740	0.9984540	0.8944252
Maximum	0.9769603	1.0000000	1.0000000	0.9823861
Missing Rate Overall	0.1493000		0.1377000	

<https://doi.org/10.1371/journal.pone.0332680.t003>

For SilicoDArT markers, the applied filter of Call rate $\geq 90\%$ and MAF of $\geq 0.5\%$ produced a set of 29,865 markers. Additionally, alignment to YP4 genome reference revealed that 22,594 SilicoDArT markers (75.7%) were successfully mapped to chromosome positions. Results of statistical analysis were very similar than obtained with SNP markers, which is shown as supplementary material (S1–S4 Figs).

Genetic Relationships and Population Structure

Clustering analysis based on the neighbor-joining algorithm resolved the evaluated accessions into two well-defined major groups, reflecting their underlying genetic relationships (Fig 9A). The first cluster comprised 284 accessions assigned to the Andean gene pool, whereas the second included 363 accessions corresponding to the Mesoamerican gene pool. This clear bifurcation underscores the strong genetic differentiation between the two domestication lineages. The first group comprised primarily accessions collected from the Peruvian Amazon, together with reference accessions included as an outgroup of Andean origin from the Peruvian highlands, the coastal region, and CIAT. The second group consisted mainly of Peruvian Amazon accessions, along with outgroup materials of Mesoamerican origin, including CIAT accessions and a reference sample from Mexico. This clear bipartite structure indicates that the first group corresponds to the Andean gene pool, whereas the second group represents the Mesoamerican gene pool, consistent with the recognized primary domestication centers of common bean.

Principal Coordinate Analysis (PCoA) further supported this genetic differentiation, with the first three axes explaining 62.62%, 8.97%, and 2.77% of the total molecular variation, respectively (Fig 9B). The first axis clearly separated the Andean and Mesoamerican gene pools, while the second and third axes captured additional within-group variation. These results confirm the strong genetic structure within the collection and highlight the coexistence of both major gene pools among common bean landraces cultivated in the Peruvian Amazon.

Population structure analysis provided additional evidence for this pattern. The Evanno method identified the strongest signal of genetic structure at $K=2$ (Fig 9D), in agreement with the cluster analysis and PCoA results. These two genetic clusters correspond to the Andean and Mesoamerican gene pools and likely reflect contrasting domestication histories, evolutionary trajectories, and adaptation processes associated with their geographic origins. At a higher level of resolution ($K=4$), two subgroups with low admixture were detected within each major group (Fig 9B), revealing finer-scale genetic differentiation. Notably, these subgroups showed a strong association with farmers' common names for the cultivars, providing valuable insight into the historical diffusion, local adaptation, and cropping practices of common bean in the Amazonian region (Fig 10). Of the Amazonia accessions, 56.1% clustered with the Mesoamerican pool and 43.9% with the Andean gene pools.

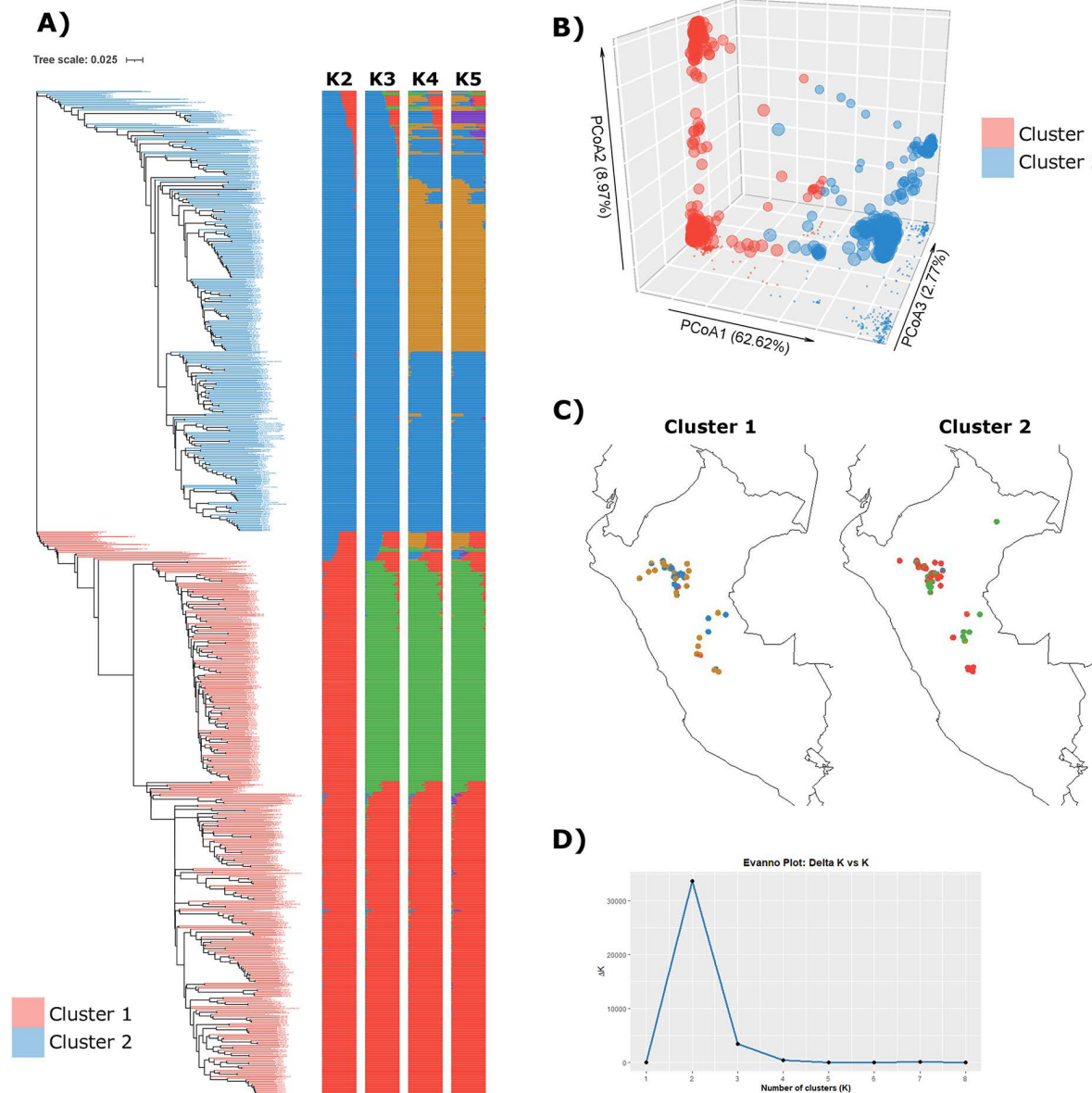


Fig 9. Genetic structure of 647 common bean accessions with 23,050 SNP markers. A) Neighbor joining tree showing two main clusters, ancestry proportions from STRUcTURE analysis from K2 to K5 (left to right). B) Principal Coordinate Analysis colored according to the two main clusters of the NJ tree. C) Pie chart of the ancestry proportions of K=4 for each accession with geographic coordinates. D) Evanno plot showing the strongest signal of population structure at K=2.

<https://doi.org/10.1371/journal.pone.0332680.g009>

Within the Andean gene pool (Group 1), two subgroups were identified and designated as subgroup 1.1 and subgroup 1.2. Subgroup 1.1 included recently collected accessions together with reference materials from the Andean highlands (including popping beans or nuñas), and coastal regions, which were incorporated to elucidate genetic relationships among Peruvian bean germplasm. This subgroup predominantly comprised accessions known locally as Canario, Norteño, Manteca, Pinto, Chuncho, 60 Días, and Caballero. In contrast, subgroup 1.2 consisted exclusively of recently collected accessions from the Peruvian Amazon and was mainly represented by landraces referred to as Ashpa,

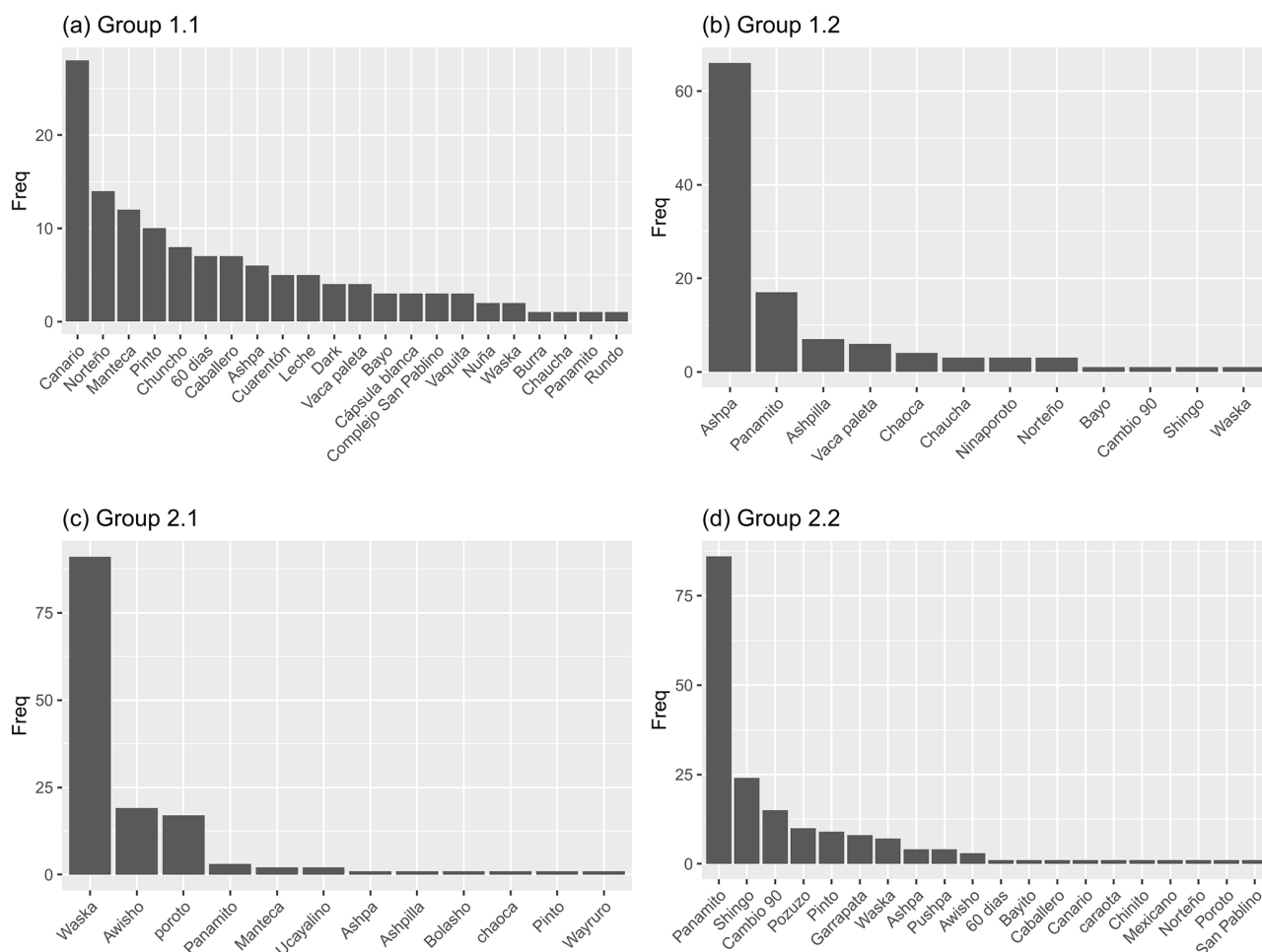


Fig 10. Number of common names distribution in each subgroup of common bean from Amazonia, Peru.

<https://doi.org/10.1371/journal.pone.0332680.g010>

Panamito, Ashpilla, and Vaca Paleta (Fig 10). The genetic distinctiveness of this subgroup suggests local differentiation and adaptation under Amazonian agroecological conditions.

Similarly, the Mesoamerican gene pool (Group 2) was subdivided into subgroup 2.1 and subgroup 2.2. Subgroup 2.1 comprised exclusively recently collected Amazonian accessions, predominantly cultivated under the common names Waska, Awisho, and Poroto. Subgroup 2.2 included both Amazonian accessions and outgroup materials of Mesoamerican origin and was mainly composed of landraces known as Panamito, Shingo, Cambio 90, Pozuzo, Pinto, and Garrapata (Fig 10). The presence of both local and reference materials within this subgroup indicates historical introductions and subsequent integration of Mesoamerican germplasm into Amazonian farming systems.

Overall, the combined clustering, PCoA, and population structure analyses reveal a complex genetic architecture of common bean landraces in the Peruvian Amazon, shaped by the coexistence of the Andean and Mesoamerican gene pools, localized diversification, and farmer-mediated selection. These findings underscore the importance of Amazonian landraces as reservoirs of genetic diversity and highlight their potential value for conservation and breeding programs targeting adaptation to diverse and changing environments.

Genetic Diversity (α and β) using “q” profiles

Using the two main groups identified in the Peruvian Amazon common bean panel, we quantified within-group (alpha, α) and between-group (beta, β) diversity across Hill numbers for $q=0$, 1, and 2 based on the 23,050 SNP markers. The results are shown in [Table 4](#).

For $q=0$ (allelic richness), α was 1.97 in Group1 and 1.95 in Group2, which is expected in biallelic SNPs, while β was 1.04, indicating that both groups harbor essentially the same set of allelic states with virtually no private alleles unique to either group. At $q=1$ (Shannon diversity), α was identical in both groups (1.31) and β increased to 1.20, which indicates that although within-group allele-frequency balance is similar, the groups begin to differ in how common those alleles are between populations. At $q=2$ (Gini–Simpson/expected heterozygosity), α was 1.22 in Group1 and 1.23 in Group2, whereas β rose to 1.4–1.5, indicating a clearer differentiation driven primarily by shifts in the frequencies of common alleles.

Taken together, these q-profiles show minimal presence/absence differentiation but progressively stronger frequency-based divergence from $q=0$ to $q=2$, indicating that the genetic structuring between groups is explained more by differences in the prevalence of shared alleles than by unique allelic repertoires.

Using the same 23,050 SNPs, we extended the q-profile analysis to four populations defined by previous cluster and STRUCTURE analysis. The results of the analysis in the four subgroups are shown in [Table 5](#).

For $q=0$, α ranged narrowly from 1.82 to 1.94 (see [Table 5](#)), indicating that all populations harbor nearly the full set of biallelic states with only minor differences in allelic richness. At $q=1$, α values differentiated the populations more clearly (Subgroup1_1 = 1.31; Subgroup1_2 = 1.25; Subgroup2_1 = 1.16; Subgroup2_2 = 1.30), which indicates varying degrees of allele frequency balance: Subgroup1_1 and Subgroup2_2 exhibit more even allele frequencies, whereas Subgroup2_1 shows greater skew toward common alleles. At $q=2$, the ranking persisted (Subgroup2_1 = 1.10; Subgroup1_2 = 1.16; Subgroup2_2 = 1.21; Subgroup1_1 = 1.22), indicating that differences among populations are driven primarily by the frequencies of common alleles.

Table 4. Hill number of α (intrapopulation) and β (interpopulation) diversity indices according to the 23,050 loci SNP and two groups of common beans from Amazonia, Peru.

Hill numbers (D)	Measures	Alpha diversity (α)		Beta diversity (β)
		Group1	Group2	Differentiation of Group 1–Group 2
$q=0$ (allele richness)	Potential richness	1.97	1.95	~1.04
$q=1$ (Shannon index)	Frequency balance	1.31	1.31	~1.2
$q=2$ (expected heterozygosity)	Effective heterozygosity (weight on common alleles)	1.22	1.23	~1.4–1.5

<https://doi.org/10.1371/journal.pone.0332680.t004>

Table 5. Hill number (D) of α (intrapopulation) and β (interpopulation) diversity indices according to the 23,050 single nucleotide polymorphism (SNP) loci and four subgroups of common beans from Amazonia, Peru.

Subgroups	D			Diversity
	$q=0$	$q=1$	$q=2$	
Subgroup1_1	1.93	1.31	1.22	α
Subgroup1_2	1.89	1.25	1.16	
Subgroup2_1	1.82	1.16	1.1	
Subgroup2_2	1.94	1.3	1.21	
Comparison subgroup1_1 vs subgroup1_2	1.06	1.24	~1.29	β
Comparison subgroup2_1 vs subgroup2_2	1.07	1.29	~1.32	

<https://doi.org/10.1371/journal.pone.0332680.t005>

Taken together, the four-way comparison shows minimal differences in allelic repertoires ($q=0$) but increasing divergence at $q=1$ and $q=2$, which indicates that genetic differentiation among these subpopulations is explained less by the presence or absence of allelic states and more by shifts in the prevalence and balance of shared alleles.

Comparing the two subgroups within each main group, we observe little presence/absence differentiation but a clear frequency-based divergence. For subgroup1_1 versus subgroup1_2, β increased from 1.06 at $q=0$ to 1.24 at $q=1$ and ~ 1.29 at $q=2$, which indicates that while both subgroups share essentially the same allelic states, they differ modestly to moderately in allele-frequency balance, with the strongest contrast driven by common alleles. A similar behavior is observed for subgroup2_1 versus subgroup2_2 (Table 5).

Population Heterozygosity

The parameters of diversity using heterozygosity, according to the 23,050 loci SNP, two groups and four subgroups of common beans are shown in Table 6.

Two-groups analysis

Group 1: where the effective sample size was ~ 240 individuals, showing 22,281 polymorphic loci (almost all, 97% of the total). The $H_o=0.041$, meaning that on average, $\sim 4.1\%$ of the loci per individual are heterozygous. Then, the $H_e/uH_e=\sim 0.148$, under Hardy-Weinberg Equilibrium (HWE) shows around $\sim 14.8\%$ of heterozygous loci would be expected. The $FIS=0.52$ (inbreeding coefficient) indicated a severe heterozygote deficit. The population is less heterozygous than expected, which reflects inbreeding, which is an expected result in a self-pollinating population, such as a common bean.

Group 2: The effective sample size was ~ 318 individuals, showing 21,846 polymorphic loci (94.8% of the total). The $H_o=0.035$, around $\sim 3.5\%$ of loci are heterozygous on average, even lower than group 1. The proportion $H_e/uH_e=\sim 0.147$ is like group 1 in expected heterozygosity, and $FIS=0.54$, also shows a strong heterozygote deficit, very comparable to group 1. Comparing both Group 1 vs. Group 2, they show very similar H_e (~ 0.15), indicating that the allele pool and frequencies are similar. H_o is lower in group 2 than in group 1 (0.035 vs. 0.041), suggesting that individuals in group 2 express less heterozygosity. The FIS is elevated in both populations (>0.5), reinforcing the idea of a generalized heterozygote deficit, which is expected for common bean populations. Finally, group2 has more monomorphic loci than group 1 (1,204 vs 769), which could reflect loss of variation or fixation at certain sites.

Four subgroups analysis

The subgroup1_1 has an effective size of ~ 136 individuals and 21,511 polymorphic loci (93% of the total). The $H_o=0.032$, quite low compared to $H_e=0.145$; and the $FIS=0.58$, reflecting a strong heterozygote deficit.

Table 6. Genetic diversity parameters within groups and subgroups with the 23,050 loci SNP of common beans from Amazonia, Peru.

Groups	n.Ind (ef.)	n.Loc	polyLoc	monoLoc	Ho ($\pm$ SD)	He ($\pm$ SD)	uHe ($\pm$ SD)	FIS ($\pm$ SD)
Group_1	~ 240	23,050	22,281	769	0.041 (0.056)	0.148 (0.145)	0.148 (0.145)	0.523 (0.364)
Group_2	~ 318	23,050	21,846	1,204	0.035 (0.051)	0.147 (0.165)	0.147 (0.165)	0.540 (0.365)
Subgroup1_1	~ 136	23,050	21,511	1,539	0.032 (0.054)	0.145 (0.148)	0.146 (0.149)	0.576 (0.384)
Subgroup1_2	~ 104	23,050	20,444	2,606	0.053 (0.068)	0.115 (0.124)	0.116 (0.125)	0.348 (0.380)
Subgroup2_1	~ 131	23,050	18,877	4,173	0.029 (0.058)	0.071 (0.114)	0.071 (0.114)	0.302 (0.380)
Subgroup2_2	~ 187	23,050	21,563	1,487	0.040 (0.055)	0.139 (0.149)	0.139 (0.149)	0.525 (0.355)

n.Ind (ef.) = effective sample size; n.Loc = number of loci; polyLoc = number of polymorphic loci; monoLoc = monomorphic loci; H_o (Observed heterozygosity; proportion of heterozygous loci per individual, average across the population); H_e (Expected heterozygosity, Nei 1978; expected heterozygosity under Hardy-Weinberg equilibrium); uH_e (unbiased H_e , corrected for sample size); FIS (Inbreeding coefficient = $1 - H_o / uH_e$, positive values indicate a deficit of heterozygotes (possible inbreeding or structure) and negative values indicate an excess).

<https://doi.org/10.1371/journal.pone.0332680.t006>

The subgroup 1_2 has an effective size of ~104 individuals and 20,444 polymorphic loci (88.7%), more monomorphic loci than subgroup1_1. The $H_o = 0.053$, higher than subgroup1_1, individuals show slightly more heterozygosity. However, the $H_e = 0.115$ was lower than subgroup1_1; and the $F_{IS} = 0.35$, still positive but lower heterozygote deficit than in subgroup1_1.

Subgroup2_1 has an effective size of ~131 individuals and 18,877 polymorphic loci (82%), the lowest of the four subgroups. Also, the $H_o = 0.029$ and $H_e = 0.071$ were the lowest of all subgroups, indicating a loss of diversity; and the $F_{IS} = 0.30$ showed a moderate heterozygote deficiency.

Subgroup2_2 has an effective size of ~187 individuals and 21,563 polymorphic loci (93.5%), comparable to subgroup1_1. The $H_o = 0.040$ showed an intermediate value between subgroup1_1 and subgroup1_2; and the $H_e = 0.139$ was intermediate too; and $F_{IS} = 0.53$, indicating a severe heterozygote deficiency.

Comparisons between subgroups within the group

Group_1: the subgroup1_1 has a higher H_e (0.145) but a lower H_o , resulting in a very high F_{IS} (0.58). The subgroup1_2 has a lower H_e (0.115) but a relatively higher H_o (0.053), which reduces the deficit of heterozygosity ($F_{IS} = 0.35$). The subgroup1_1 appears more structured as a self-pollinating population, while subgroup1_2 maintains more heterozygotes despite lower overall diversity.

Group 2: the subgroup2_1 has the lowest diversity among the four subgroups ($H_o = 0.029$; $H_e = 0.071$). However, the subgroup2_2 recovers diversity ($H_e = 0.139$), but its F_{IS} is still high (0.53). Within Group_2, there is a clear inequality; subgroup2_1 present loss of diversity; and subgroup2_2 is more diverse but lacks heterozygotes.

In all four subgroups, there is a clear disconnection between H_o and H_e ; all subgroups have fewer heterozygotes than expected under the Hardy–Weinberg equilibrium. Also, the positive and high F_{IS} (0.30–0.58) in all subgroups suggest inbreeding, internal structure favoring homozygotes, which is expected in common bean.

Discussion

Morphologic variability

The morphological diversity observed in terms of type of plant, flowering time, color, size, form of pods and seeds confirms that Amazonian *P. vulgaris* represents an important genetic reservoir. The presence of distinct genetic clusters suggests that both environmental heterogeneity and farmer practices shape phenotypic variability. Similar levels of morphological diversity have been reported in Brazil [42], Mexico [43], Spain [44], Nigeria [45], Pakistan [46], Ethiopia [47]. In Peru, however, most previous studies have focused on highland germplasm, using small sample size [21–23,48] and lately sampling from Amazonia [25]; which are not representative to allow a comprehensive estimation of common bean variability. Our results show that much of the morphological diversity is conserved on farms, where farmers save, select, and replant seeds year after year, ensuring both autonomy and household income [49,50]. In addition, frequent seed acquisition by farmers from traditional markets further promotes gene flow and admixture, particularly relevant under tropical conditions where high temperature and humidity shorten seed longevity, and its storage challenges; causes rapid loss of viability. Together, these practices contribute to the maintenance and renewal of agrobiodiversity in the Peruvian Amazon. It is relevant to highlight the farmer behaviors, who collect new material and sow all together with their own local landraces. Then, the agroecological mosaic, coupled with traditional knowledge, has led to the maintenance of a broad range of landraces with varying morphological traits. A large portion of this agrobiodiversity is still conserved *in situ*, within Indigenous and local farming systems. Farmers contribute to the conservation and evolution of these cultivars by using and exchanging seeds year by year [51,52]. Additionally, field observations and interviews revealed that many farmers regularly acquire seeds from traditional markets to replace planting stock each season. This behavior is likely influenced by tropical environmental constraints—such as high temperatures and humidity—which accelerate seed aging and reduce

viability during storage. Thus, constant renewal of planting material becomes a necessary strategy. The morphological diversity recorded—particularly in traits like growth habit, seed color, and pod characteristics—supports the idea of local adaptation and cultural preferences influencing varietal selection. Morphological characterization, though often considered basic, remains a valuable and accessible tool for identifying distinct genotypes and understanding their agronomic potential and farmer preferences.

Molecular diversity

SNP-based clustering and Principal Coordinate Analysis (PCoA) separated the accessions into two distinct groups corresponding to the Andean and Mesoamerican gene pools, comprising 284 and 363 accessions, respectively (Fig 9). This pattern is consistent with the well-established genetic differentiation between the two major gene pools of *Phaseolus vulgaris* and confirms that genome-wide SNP markers effectively capture the species' population structure. The presence of Amazonian accessions in both gene pools indicates that the region harbors genetic diversity representative of the two major lineages of common bean. However, this tendency was largest in others common bean germplasm genetic diversity studies, e.g., in Brazil was reported, more than 70% of Mesoamerican gene pool, [4, 53]; and in Ethiopian germplasm, the 284 common bean accessions were roughly classified into Andean and Mesoamerican gene pools, assigning 188 (65.05%) accessions in the Mesoamerican group, while the remaining 96 (33.22%) accessions were classified in the Andean gene pool [28]. The population structure analysis (based on $K=2$ groups) showed that the common bean accessions from the Peruvian Amazon were distributed between two gene pools, Andean and Mesoamerican, with a low degree of admixture.

This structure is consistent with findings from USDA core collections [54], where PCoA and NJ trees confirmed the Andean–Mesoamerican split and further subdivisions into ecogeographic races. These results indicate that Amazonian germplasm is integrated into the global diversity framework of *P. vulgaris* while retaining unique local substructure. However, in Brazilian germplasm genetic diversity research in common bean, more diversity was obtained in the Mesoamerican gene pool using GBS analysis [53]. Overall, the genetic variation detected through DarTseq SNP analysis showed that the clustering agrees with that obtained by the PCoA and population structure analysis.

The high genetic variation detected through DarTseq SNP analysis complements the phenotypic findings and highlights the importance of both *in situ* and *ex situ* conservation. Many of these accessions, particularly those maintained only on farm, are at risk of being lost due to changes in land use, market pressures, climate change, or generational shifts in farming practices. Comprehensive genetic characterization enables the identification of unique and underrepresented alleles, which can be prioritized for conservation and potentially incorporated into breeding programs aimed at improving resilience to biotic and abiotic stresses.

Alpha (α) and Beta (β) Genetic Diversity

The Peruvian Amazonia common beans, according to the alpha-diversity estimates, showed similar internal variability in both gene pools, with most loci being biallelic and rare alleles contributing little or marginally to overall diversity. Beta-diversity revealed that Andean and Mesoamerican groups largely share the same alleles but differ in allele frequencies, reflecting the combined effects of local selection, drift, and demographic histories.

The four subgroups identified at $K=4$ maintained nearly complete allelic richness, with differentiation arising mainly from differences in allele frequencies rather than unique variants. This pattern is consistent with historical farmer migration from the Andes into Amazonian areas, where likely Andean beans were introduced and later mixed with locally adapted materials, followed by continuous selection and cultivation under tropical conditions. This result is similar to what was found by Blair et al. [55] on diversification and population structure in common beans, suggesting geographic isolation, founder effects, or natural selection and introgression were involved in creating the diversity of cultivated beans.

Population Heterozygosity

Despite the high proportion of polymorphic *loci*, observed heterozygosity ($H_o \approx 0.035\text{--}0.041$) was markedly lower than expected ($H_e \approx 0.147\text{--}0.148$). Positive and high inbreeding coefficients ($F_{IS} = 0.30\text{--}0.58$) across all groups and subgroups indicate strong homozygosity, which is expected in a predominantly self-pollinating crop such as common bean. Similar patterns of heterozygote deficit have been reported in Brazil [42] and Ethiopia [28].

In both morphological and molecular characterization, no duplicates were found in the material under study. This result is like that found when evaluating 67 accessions based on morphological traits and molecular markers from Brazil [42], which suggests the importance of keeping all accessions in the germplasm bank since they represent a valuable source of genetic diversity. Importantly, this variability has remained largely unexploited in breeding programs, which have focused on disease resistance, lagging in obtaining high-yielding and high-quality cultivars in major production regions [56,57]. Incorporating Amazonian diversity into breeding pipelines could broaden the adaptive base of the common bean. Hence, the conservation and systematic use of this germplasm are of strategic importance for Peru's national genetic resources and future crop improvement.

Conclusion

This study addresses the limited characterization of genetic diversity in common landraces from the Peruvian Amazon, a region that has been largely overlooked in favor of highland and coastal germplasm. Our results show that these landraces harbor substantial morphological and genomic diversity, confirming the Amazon as an important reservoir of agrobiodiversity. Both morphological descriptors and SNP-based analyses consistently identified two major groups corresponding to the Andean and Mesoamerican gene pools, each further subdivided into distinct subgroups that include unique Amazonian accessions. High allelic richness combined with low observed heterozygosity was consistent with the predominantly self-pollinating nature of the species and the associated high levels of inbreeding, while genetic variation was primarily explained by differences in allele frequencies rather than the presence of unique alleles, indicating similar levels of internal diversity across groups. The absence of redundant accessions highlights the distinct genetic contribution of each sampled landrace. The combined use of morphological traits and DArTseq SNP markers proved effective for detecting population structure and diversity patterns, supporting their complementary role in germplasm characterization. Given that much of this diversity is maintained in situ by local farmers, these findings highlight the importance of strengthening both in situ and ex situ conservation strategies. Overall, this work provides a baseline for future research and supports the use of Amazonian common bean landraces as a valuable resource for breeding programs aimed at improving adaptation to tropical environments and resilience to changing climatic conditions.

Supporting information

S1 Fig. Histogram distribution of SilicoDArT marker parameters. A) Callrate, B) OneRatio, C) Polymorphic Index Content (PIC), and E) minimum allele frequency (MAF) on the initial filtered set 29865 SNP markers.
(TIFF)

S2 Fig. Dendrogram of 647 accessions of common bean, according to the 29865 SilicoDArT markers.
(TIFF)

S3 Fig. Principal coordinate analysis (PCoA) of 647 accessions of common bean according to SilicoDArT markers.
(TIFF)

S4 Fig. Population structure of 647 accessions of common bean with SNP-DArTseq markers. K2 to K5 (from left to right).
(TIFF)

S1 File. Supplementary_data. (DOCX)

Acknowledgments

The authors are thankful to SAGA-Genetic Analysis Service for Agriculture and CIMMYT (Mexico) for their support for genotyping. Also, we thank local farmers for their cooperation and the research institutions (UNALM, UNSM, UNU) that supported this study.

Author contributions

Conceptualization: Jorge Tobaru, Marvin Barrera-Lozano, Carlos D. Vecco-Giove, Mack Pinchi, Raul Blas.

Data curation: Ronald Robles, David Saravia, Cesar Petrolí.

Formal analysis: Ronald Robles, David Saravia, Cesar Petrolí.

Funding acquisition: Raul Blas.

Investigation: Jorge Tobaru, Jorge Peláez, Yesenia Vargas, Jorge Chuquillanqui.

Methodology: Jorge Tobaru, Marvin Barrera-Lozano, Carlos D. Vecco-Giove, Mack Pinchi, Raul Blas.

Supervision: Rodomiro Ortiz, Raul Blas.

Visualization: Ronald Robles, David Saravia, Cesar Petrolí.

Writing – original draft: Jorge Tobaru, Marvin Barrera-Lozano, Carlos D. Vecco-Giove, Mack Pinchi, Ronald Robles, David Saravia, Cesar Petrolí.

Writing – review & editing: Julian Soto-Torres, Jorge D. Etchevers, Rodomiro Ortiz.

References

1. Wang Y, Xiaopeng H, Chunhai C, Haigang W, Peng G, Xukui Y. Telomere-to-telomere genome of common bean (*Phaseolus vulgaris* L.) YP4. *Giga-Science*. 2025;14:giaf001. <https://doi.org/10.1093/gigascience/giaf001>
2. Bitocchi E, Nanni L, Bellucci E, Rossi M, Giardini A, Zeuli PS, et al. Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data. *Proc Natl Acad Sci U S A*. 2012;109(14):E788–96. <https://doi.org/10.1073/pnas.1108973109> PMID: 22393017
3. Kwak M, Kami JA, Gepts P. The Putative Mesoamerican Domestication Center of *Phaseolus vulgaris* Is Located in the Lerma–Santiago Basin of Mexico. *Crop Science*. 2009;49(2):554–63. <https://doi.org/10.2135/cropsci2008.07.0421>
4. Bellucci E, Bitocchi E, Rau D, Rodriguez M. Genomics of origin, domestication and evolution of *Phaseolus vulgaris*. In: Tuberosa Rde, Graner A, Frison E, editors. Dordrecht, The Netherlands: Springer. 2014. p. 483–507. https://doi.org/10.1007/978-94-007-7572-5_20
5. Frascarelli G, Galise TR, D'Agostino N, Cafasso D, Cozzolino S, Cortinovis G, et al. The evolutionary history of the common bean (*Phaseolus vulgaris*) revealed by chloroplast and nuclear genomes analysis. *Theor Appl Genet*. 2025;138(2):47. <https://doi.org/10.1007/s00122-025-04832-z> PMID: 39920343
6. Singh SP. Broadening the genetic base of common bean cultivars. *Crop Science*. 2001; 41: 1659–75. <https://doi.org/10.2135/cropsci2001.1659>
7. Gepts P. Origin and Evolution of Common Bean: Past Events and Recent Trends. *HortSci*. 1998;33(7):1124–30. <https://doi.org/10.21273/hortsci.33.7.1124>
8. Schmutz J, McClean PE, Mamidi S, Wu GA, Cannon SB, Grimwood J, et al. A reference genome for common bean and genome-wide analysis of dual domestications. *Nat Genet*. 2014;46(7):707–13. <https://doi.org/10.1038/ng.3008> PMID: 24908249
9. Angioi SA, Rau D, Attene G, Nanni L, Bellucci E, Logozzo G. Beans in Europe: origin and structure of the European landraces. *Theoretical and Applied Genetics*. 2010;121:829–43.
10. Paredes C M, Becerra V V, Tay U J. Inorganic Nutritional Composition of Common Bean (*Phaseolus vulgaris* L.) Genotypes Race Chile. *Chilean J Agric Res*. 2009;69(4). <https://doi.org/10.4067/s0718-58392009000400002>
11. Suárez-Martínez SE, Ferriz-Martínez RA, Campos-Vega R, Elton-Puente JE, de la Torre Carbot K, García-Gasca T. Bean seeds: leading nutraceutical source for human health. *CyTA - Journal of Food*. 2015;14(1):131–7. <https://doi.org/10.1080/19476337.2015.1063548>
12. Jensen ES, Peoples MB, Boddey RM, Gresshoff PM, Hauggaard-Nielsen H, J.R. Alves B, et al. Legumes for mitigation of climate change and the provision of feedstock for biofuels and biorefineries. A review. *Agron Sustain Dev*. 2011;32(2):329–64. <https://doi.org/10.1007/s13593-011-0056-7>

13. United Nations. International Year of Pulses, 2016. 2014. <https://docs.un.org/en/A/RES/68/231>
14. Saburido M, Herrera A. El frijol en la era genómica. *Revista Digital Universitaria*. 2015;2:16.
15. Calero-Hurtado A, Quintero-Rodríguez E, Olivera-Viciedo D, Pérez-Díaz Y, Castro-Lizazo I, Jiménez J. Respuesta de dos cultivares de frijol común a la aplicación foliar de microorganismos eficientes. *Cultivos Tropicales*. 2018;39:5–10.
16. Food and Agriculture Organization of the United Nations. Legumbres. Pequeñas semillas, grandes soluciones. Rome, Italy: Food and Agriculture Organization of the United Nations. 2018.
17. Ministerio de Agricultura y Desarrollo Rural. El agro en cifras. Ministerio de Agricultura y Desarrollo Rural. 2024. https://siea.midagri.gob.pe/portal/phocadownload/datos_estadisticas/mensual/Agro/2024/Agro_en_cifras%2006%202024.pdf
18. Genesys. Genesys Global Portal on Plant Genetic Resources. www.genesys-prg.org. 2021.
19. I. N. I. A. Catálogo de ñuña en el banco de germoplasma del INIA. <https://hdl.handle.net/20.500.12955/2112>. 2021.
20. I. N. I. A. Catálogo de frijol en regiones andinas del banco de germoplasma del INIA. 2023. <https://hdl.handle.net/20.500.12955/1404>
21. Balarezo JC. Evaluación agromorfológica y caracterización molecular de la ñuña (*Phaseolus vulgaris* L.). *Idesia (Arica)*. 2009;27:29–40. <https://doi.org/10.4067/S0718-34292009000100005>
22. Marmolejo K. Variabilidad genética del frijol común tipo ñuña (*Phaseolus vulgaris* L.) en las localidades de Carhuaz y Chiquián, Ancash. *Universidad Nacional Agraria La Molina*. 2018. <http://repositorio.lamolina.edu.pe/handle/UNALM/3398>
23. Amaringo S. Caracterización fenotípica de 5 ecotipos de frijol común (*Phaseolus vulgaris* L.) en el fundo Aucaloma – Lamas. Tarapoto: Universidad Nacional de San Martín. 2015. <http://hdl.handle.net/11458/635>
24. Collado LA. Diversidad cultivada y socio-cultural en la Amazonia central del Perú. *Universidad Nacional Agraria de la Selva*. 2002. <http://repositorio.unas.edu.pe/handle/UNAS/900>
25. Vásquez J, Vilca-Valqui NC, Malqui R, Fernández E, Duarez E, Ayala R. Agromorphological characterization of *Phaseolus* spp., accessions in Amazonas region, Peru. *Bioagro*. 2024;36:129–42. <https://doi.org/10.51372/bioagro362.1>
26. Seo E, Kim K, Jun T-H, Choi J, Kim S-H, Muñoz-Amatriain M, et al. Population Structure and Genetic Diversity in Korean Cowpea Germplasm Based on SNP Markers. *Plants (Basel)*. 2020;9(9):1190. <https://doi.org/10.3390/plants9091190> PMID: 32932572
27. Verma S, Gupta S, Bandhiwal N, Kumar T, Bharadwaj C, Bhatia S. High-density linkage map construction and mapping of seed trait QTLs in chick-pea (*Cicer arietinum* L.) using Genotyping-by-Sequencing (GBS). *Sci Rep*. 2015;5:17512. <https://doi.org/10.1038/srep17512> PMID: 26631981
28. Gelaw YM, Eleblu JSY, Ofori K, Fenta BA, Mukankusi C, Emam EA, et al. High-density DARTSeq SNP markers revealed wide genetic diversity and structured population in common bean (*Phaseolus vulgaris* L.) germplasm in Ethiopia. *Mol Biol Rep*. 2023;50(8):6739–51. <https://doi.org/10.1007/s11033-023-08498-y> PMID: 37389701
29. International Board for Plant Genetic Resources IBPGR. *Phaseolus vulgaris* descriptors. Rome, Italy: International Board for Plant Genetic Resources. 1982.
30. Team RC. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. 2025.
31. Posit team. RStudio: Integrated Development Environment for R. Boston, MA: Posit Software. 2025.
32. Lê S, Josse J, Husson F. FactoMineR: An R Package for Multivariate Analysis. *J Stat Soft*. 2008;25(1). <https://doi.org/10.18637/jss.v025.i01>
33. Kassambara A, Mundt F. Factoextra: extract and visualize the results of multivariate data analyses. 2020. <https://doi.org/10.32614/CRAN.package.factoextra>
34. Wickham H. ggplot2: Elegant graphics for data analysis. New York: Springer-Verlag. 2016.
35. Liaw A, Wiener M. Classification and regression by randomForest. *R News*. 2002;2:18–22.
36. Sansaloni C, Petrolí C, Jaccoud D, Carling J, Detering F, Grattapaglia D, et al. Diversity Arrays Technology (DART) and next-generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of Eucalyptus. *BMC Proc*. 2011;5(S7). <https://doi.org/10.1186/1753-6561-5-s7-p54>
37. Kilian A, Wenzl P, Huttner E, Carling J, Xia L, Blois H, et al. Diversity arrays technology: a generic genome profiling technology on open platforms. *Methods Mol Biol*. 2012;888:67–89. https://doi.org/10.1007/978-1-61779-870-2_5 PMID: 22665276
38. Petrolí C, Kilian A. Introduction to the DARTseq genotyping method and its data outputs. 2019. <https://hdl.handle.net/11529/10548358/>
39. Gruber B, Unmack PJ, Berry OF, Georges A. dart: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Mol Ecol Resour*. 2018;18(3):691–9. <https://doi.org/10.1111/1755-0998.12745> PMID: 29266847
40. Perrier X, Jacquemoud-Collet JP. DARwin software: dissimilarity analysis and representation for Windows. <https://darwin.cirad.fr/>. 2006.
41. Raj A, Stephens M, Pritchard JK. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*. 2014;197(2):573–89. <https://doi.org/10.1534/genetics.114.164350> PMID: 24700103
42. de Paula E, Almeida RN de, Santos T de O, Souza Neto JD de, Riva-Souza EM, Posse SCP, et al. Genetic Diversity of Common Bean (*Phaseolus vulgaris* L.) Landraces Based on Morphological Traits and Molecular Markers. *Plants (Basel)*. 2024;13(18):2584. <https://doi.org/10.3390/plants13182584> PMID: 39339559

43. García-Narváez AL, Hernández-Delgado S, Hernández-Delgado JL, Mayek-Pérez N. Morphological and agronomic variability of bean germplasm cultivated in Oaxaca, México. *Revista Bio Ciencias*. 2020;7.
44. Arteaga S, Yabor L, Torres J, Solbes E, Muñoz E, Díez MJ. Morphological and agronomic characterization of Spanish landraces of *Phaseolus vulgaris* L. *Agriculture*. 2019;9(7):149.
45. Kouam EB, Kanga-Fotso AM-A, Anoumaa M. Exploring agro-morphological profiles of *Phaseolus vulgaris* germplasm shows manifest diversity and opportunities for genetic improvement. *Journal of Agriculture and Food Research*. 2023;14:100772. <https://doi.org/10.1016/j.jafr.2023.100772>
46. Nawaz I, Zeb T, Zeb BS, Sherani J. Evaluation of Common Beans (*Phaseolus vulgaris* L.) Land Races based on Qualitative Traits Derived from Himalayan, Andean and Mesoamerican regions. *SJA*. 2021;38(1). <https://doi.org/10.17582/journal.sja/2022/38.1.275.286>
47. Tigist SG, Sibiya J, Amelework A, Keneni G. Agromorphological and Physiological Performance of Ethiopian Common Bean (*Phaseolus vulgaris* L.) Genotypes under Different Agroecological Conditions. *Plants (Basel)*. 2023;12(12):2342. <https://doi.org/10.3390/plants12122342> PMID: [37375967](https://pubmed.ncbi.nlm.nih.gov/37375967/)
48. Castro R, Quispe D. Caracterización morfológica de genotipos de frijol (*Phaseolus vulgaris* L.) en San Juan de Uchubamba-Jauja, La Merced Perú. Universidad Nacional Daniel Alcides Carrión. 2017.
49. Figueiredo Menezes Cavalcanti T, Pombo Sudré C, Wesley da Silva Corrêa J, dos Santos Bento C, Knoblauch Viegas de Andrade E, Kazue Nakamura Fukuiji K, et al. Custodians of common bean diversity in Rio de Janeiro state, Brazil: revealing their socioeconomic and environmental profile. *Agroecology and Sustainable Food Systems*. 2021;45:1165–88. <https://doi.org/10.1080/21683565.2021.1888186>
50. de Avila JET, Coelho FC, Sardo PML, Haddade IR, Moonen A-C. In situ conservation of agrobiodiversity carried out by traditional family farmers in a mountainous region of Brazil. *Agroecology and Sustainable Food Systems*. 2024;49(6):797–820. <https://doi.org/10.1080/21683565.2024.2445729>
51. Castañeda-Álvarez NP, de Haan S, Juárez H, Khoury CK, Achicanoy HA, Sosa CC, et al. Ex situ conservation priorities for the wild relatives of potato (*Solanum* L. Section *petota*). *PLoS One*. 2015;10(4):e0122599. <https://doi.org/10.1371/journal.pone.0122599> PMID: [25923711](https://pubmed.ncbi.nlm.nih.gov/25923711/)
52. Bellon MR, Hellin J. Planting Hybrids, Keeping Landraces: Agricultural Modernization and Tradition Among Small-Scale Maize Farmers in Chiapas, Mexico. *World Development*. 2011;39(8):1434–43. <https://doi.org/10.1016/j.worlddev.2010.12.010>
53. Elias JCF, Gonçalves-Vidigal MC, Ariani A, Valentini G, Martiniano-Souza M da C, Vaz Bisneta M, et al. Genome-Environment Association Analysis for Bio-Climatic Variables in Common Bean (*Phaseolus vulgaris* L.) from Brazil. *Plants (Basel)*. 2021;10(8):1572. <https://doi.org/10.3390/plants10081572> PMID: [34451617](https://pubmed.ncbi.nlm.nih.gov/34451617/)
54. Kuzay S, Hamilton-Conaty P, Palkovic A, Gepts P. Is the USDA core collection of common bean representative of genetic diversity of the species, as assessed by SNP diversity?. *Crop Science*. 2020;60(3):1398–414. <https://doi.org/10.1002/csc2.20032>
55. Blair MW, Soler A, Cortés AJ. Diversification and population structure in common beans (*Phaseolus vulgaris* L.). *PLoS One*. 2012;7(11):e49488. <https://doi.org/10.1371/journal.pone.0049488> PMID: [23145179](https://pubmed.ncbi.nlm.nih.gov/23145179/)
56. Singh SP, Schwartz HF. Breeding Common Bean for Resistance to Diseases: A Review. *Crop Science*. 2010;50(6):2199–223. <https://doi.org/10.2135/cropsci2009.03.0163>
57. Cichy KA, Porch TG, Beaver JS, Cregan P, Fourie D, Glahn RP, et al. A *Phaseolus vulgaris* Diversity Panel for Andean Bean Improvement. *Crop Science*. 2015;55(5):2149–60. <https://doi.org/10.2135/cropsci2014.09.0653>