**BIODIVERSITAS**
Volume 27, Number 5, May 2026
Pages: 15

ISSN: 1412-033X
E-ISSN: 2085-4722
https://doi.org/10.13057/biodiv/d270511

# Morphometric diversity of wild poisonous arbila (*Phaseolus lunatus*) germplasm from Timor and Sumba Islands, Indonesia

**AGNES V. SIMAMORA[1,♥], LINCE MUKKUN[1], MAYAVIRA V. HAHULY[1], PETRONELLA S. NENOTEK[1], YOSEP SERAN MAU[1], NORMAN RIWU KAHO[2], YOSEPH NAHAK SERAN[3], ENOS TANGKE ARUNG[4], EVERT Y. HOSANG[5]**

[1]Department of Agrotechnology, Faculty of Agriculture, Universitas Nusa Cendana. Jl. Adisucipto, Kupang 85001, East Nusa Tenggara, Indonesia. Tel.: +62-380-881580, ♥email: asimamora@staf.undana.ac.id
[2]Department of Forestry, Faculty of Agriculture, Universitas Nusa Cendana. Jl. Adisucipto, Penfui, Kupang 85148, East Nusa Tenggara, Indonesia
[3]Department of Biological Education, Faculty of Education and Teacher Training, Universitas Timor. Jl. El Tari Km. 09, Timor Tengah Utara 85613, East Nusa Tenggara, Indonesia
[4]Department of Forest Products, Faculty of Forestry, Universitas Mulawarman. Jl. Ki Hajar Dewantara, Samarinda 75123, East Kalimantan, Indonesia
[5]Research Center for Food Crops, National Research and Innovation Agency. Jl. Raya Jakarta-Bogor Km. 46, Bogor 16912, West Java, Indonesia

Manuscript received: 7 February 2026. Revision accepted: 6 May 2026.

**Abstract.** *Simamora AV, Mukkun L, Hahuly MV, Nenotek PS, Mau YS, Kaho NR, Seran YN, Arung ET, Hosang EY. 2026. Morphometric diversity of wild poisonous arbila (*Phaseolus lunatus*) germplasm from Timor and Sumba Islands, Indonesia. Biodiversitas 27 (5): d270511. https://doi.org/10.13057/biodiv/d270511.* Poisonous arbila (*Phaseolus lunatus*) is an underutilized legume traditionally consumed after detoxification by dryland communities in East Nusa Tenggara, Indonesia. Despite its importance as an emergency food resource, information on its phenotypic variation remains limited. This study aimed to characterize seed morphometric variation in wild and semi-managed arbila germplasm collected from six locations across Timor and Sumba Islands. A total of 108 accessions were evaluated using twelve seed traits, including eight qualitative characters (seed shape, seed coat color and pattern, hilum shape, testa texture, and eye color) and four quantitative characters (seed length, width, thickness, and 100-seed weight). Qualitative traits showed high polymorphism, particularly in seed coat pigmentation and patterning, while quantitative traits exhibited wide ranges (seed length: 10.33-23.70 mm; 100-seed weight: 25.40-56.50 g). Multivariate analyses (principal component analysis and hierarchical clustering) indicated that the first four principal components explained 91.6% of the total variance, with seed size–related traits dominating the primary axis (PC1 = 65.2%). Clustering results identified one major group comprising most accessions and a small number of divergent accessions, consistent with a continuous phenotypic gradient rather than discrete group separation. These results highlight considerable morphometric diversity in arbila and indicate that phenotypic differentiation is primarily structured along this gradient. The findings provide a useful basis for germplasm conservation and highlight the potential of arbila for selection, breeding, and future studies on domestication and genetic diversity.

**Keywords:** Dryland ecosystem, multivariate analysis, seed coat patterning, seed morphology, underutilized legume

**Abbreviations:** PCA: Principal Component Analysis, UPGMA: Unweighted Pair Group Method with Arithmetic Mean, PC1-PC4: Principal Components 1-4

## INTRODUCTION

Agroecological conditions in East Nusa Tenggara (ENT) Province, Indonesia, especially Timor and Sumba Islands, are characterized by irregular rainfall, prolonged dry seasons, and low soil fertility, all of which constrain agricultural productivity in predominantly rainfed farming systems. These environmental limitations contribute to recurrent seasonal food shortages and increase the vulnerability of rural communities to crop failure and food insecurity. Under such conditions, local households have traditionally relied on wild and semi-domesticated plant species as supplementary food resources, particularly during periods of scarcity locally referred to as *paceklik*. Beyond their role as emergency foods, these species are embedded within local knowledge systems that support household resilience under variable dryland environments (Riptanti et al. 2019; Antriyandarti et al. 2024).

One plant species traditionally utilized within this context is poisonous arbila (*Phaseolus lunatus* L.), internationally known as lima bean and locally referred to as arbila racun due to the presence of toxic cyanogenic glucosides, mainly linamarin and lotaustralin, in its seeds. These compounds can release hydrogen cyanide (HCN) upon hydrolysis and may pose health risks if consumed without proper processing (Ambika et al. 2022; Zhang 2024). Despite its toxicity, arbila has long been consumed in ENT through indigenous detoxification methods, particularly repeated boiling accompanied by replacement of cooking water. After adequate processing, the seeds provide carbohydrates, plant-based protein, and essential micronutrients, highlighting their value as an alternative food source during staple scarcity (Puspita et al. 2017; Naisali et al. 2023).

Although arbila continues to play a role in local food systems, scientific information regarding this species in eastern Indonesia remains limited. In particular, there is no systematic characterization of seed morphological diversity among arbila populations distributed across Timor and Sumba Islands, despite their broad ecological distribution and continued utilization by local communities. Previous studies in Indonesia have focused mainly on utilization aspects, ethnobotanical perspectives, or general species descriptions, whereas patterns of intraspecific variation in seed traits remain poorly documented. This contrasts with research conducted in Latin America and Africa, where *P. lunatus* has been extensively studied in relation to domestication history, morphological diversity, genetic structure, and environmental adaptation (Chacón-Sánchez and Martínez-Castillo 2017; Souza et al. 2023; Barros et al. 2024).

In ENT, arbila populations persist mainly as wild or semi-managed stands across upland and lowland environments with minimal agronomic input and limited human selection. Under these conditions, variation in seed morphology may reflect environmental heterogeneity, phenotypic plasticity, and low-intensity human-mediated selection. However, it remains unclear whether quantitative seed size traits vary consistently across agroecological zones and islands, and whether qualitative traits, such as seed coat color and patterning, show spatially structured distributions. Addressing these questions is important for understanding morphological diversity within local arbila germplasm and for supporting future conservation and germplasms utilization efforts.

Morphological characterization represents an important first step in evaluating plant genetic resources, particularly in regions where molecular approaches are not yet routinely accessible (Purwanti and Fauzi 2019; Fagbédji et al. 2023). Seed-related traits are especially informative in legumes because they are relatively stable, easily observed, and often strongly influenced by genetic factors. In *Phaseolus* species, seed morphology has long been used to differentiate populations, identify landraces, and examine relationships among accessions (Bria et al. 2019; Purwanti and Fauzi 2019). For underutilized crops such as arbila, analyses integrating qualitative and quantitative seed traits provide a practical approach for exploring intraspecific diversity and identifying valuable germplasm.

To address these gaps, the present study was conducted to characterize and analyze seed morphometric diversity of arbila collected from six ecologically distinct locations across Timor and Sumba Islands. Based on the predominantly wild or semi-managed status of arbila populations, we hypothesized that seed morphometric traits exhibit structured variation along phenotypic gradients, with quantitative seed size traits contributing primarily to the dominant axes of differentiation, whereas qualitative traits contribute mainly to secondary variation without forming discrete phenotypic groups. Specifically, this study addressed two questions: (i) how quantitative seed size traits vary among arbila populations across islands and agroecological settings, and (ii) whether qualitative seed traits, including seed coat color and patterning, show location-specific distribution patterns. Multivariate analyses were employed to summarize phenotypic structure and establish baseline information supporting future conservation, breeding, and domestication studies of this underutilized legume.

## MATERIALS AND METHODS

### Sampling locations

Seed sampling was conducted from May to July 2025 across six representative dryland locations in East Nusa Tenggara Province, Indonesia (Figure 1). The study sites were selected to represent major agroecological zones on Timor Island and Sumba Island, and each site corresponds to a distinct village-level sampling location, defined within a clear administrative hierarchy (village–sub-district–district–island).

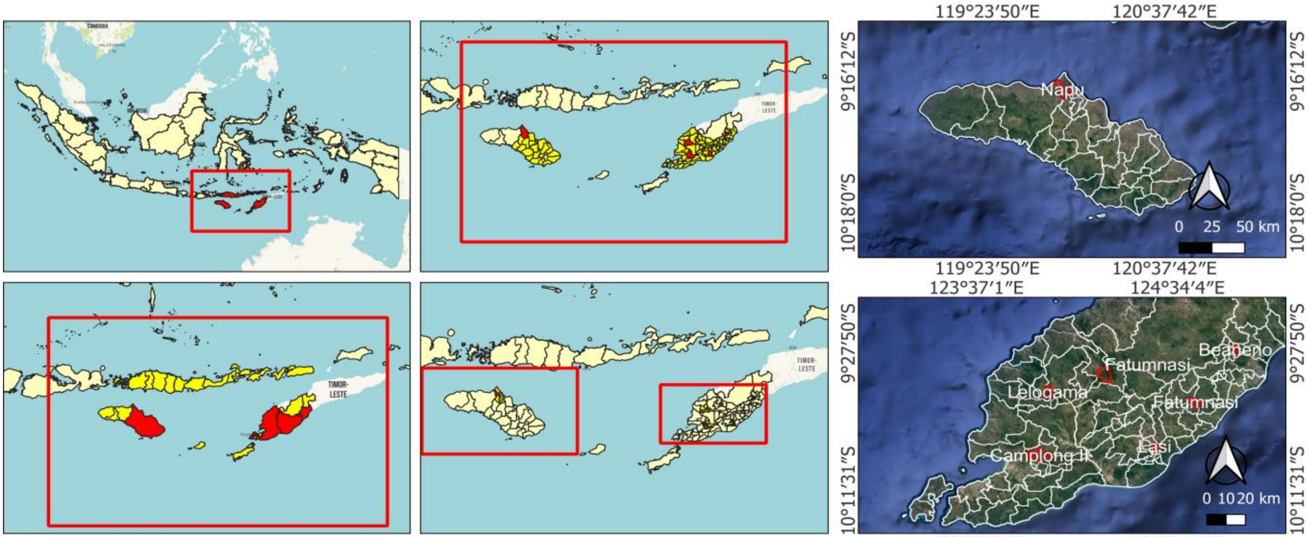


Figure 1. Sampling locations of poisonous arbila (*Phaseolus lunatus*) accessions across Timor and Sumba Islands, East Nusa Tenggara Province, Indonesia

**Table 1.** Long-term village-level climatological and topographic characteristics (1980-2025) of arbila sampling sites in East Nusa Tenggara, Indonesia

| Study site | Administrative hierarchy | Elevation (m asl.) | Temp (°C) | Precipitation (mm) | RH (%) |
|---|---|---|---|---|---|
| Lelogama Village (AF) | South Amfoang Sub-district, Kupang District, Timor Island | 889 | 24.8 | 1,556 | 77.4 |
| Camplong 2 Village (CP) | Fatuleu Sub-district, Kupang District, Timor Island | 89 | 25.6 | 1,452 | 75.2 |
| Fatumnasi Village (FT) | Fatumnasi Sub-district, South Central Timor District, Timor Island | 1,578 | 21.8 | 1,852 | 82.1 |
| Lasi Village (LS) | Kuanfatu Sub-district, South Central Timor District, Timor Island | 714 | 24.2 | 1,648 | 78.5 |
| Beaneno Village (MK) | Sasitamean Sub-district, Malaka District, Timor Island | 219 | 25.4 | 1,478 | 75.6 |
| Napu Village (SB) | Haharu Sub-district, East Sumba District, Sumba Island | 177 | 25.5 | 2,012 | 79.3 |

Note: Elevation values represent the mean village-level elevation extracted from the 30-m Shuttle Radar Topography Mission (SRTM) digital elevation model. Climatic variables (temperature, precipitation, and relative humidity) represent long-term annual means (1980-2025) derived from the ERA5-Land reanalysis dataset (ECMWF). Climatic variables are provided for site characterization only and were not used in trait-environment association analyses

The six sampling locations were Lelogama Village, South Amfoang Sub-district, Kupang District (AF); Camplong 2 Village, Fatuleu Sub-district, Kupang District (CP); Fatumnasi Village, Fatumnasi Sub-district, South Central Timor District (FT); Lasi Village, Kuanfatu Sub-district, South Central Timor District (LS); Beaneno Village, Sasitamean Sub-district, Malaka District (MK); and Napu Village, Haharu Sub-district, East Sumba District (SB). All locations are situated on Timor Island, except East Sumba, which is located on Sumba Island. For analytical consistency, each village was considered a single sampling unit.

To provide environmental context and improve reproducibility, long-term climatic and topographic characteristics of each study site were derived using the ERA5-Land reanalysis dataset (1980-2025) from the European Center for Medium-Range Weather Forecasts (ECMWF), processed via the Google Earth Engine (GEE) platform. Mean annual air temperature, precipitation, and relative humidity were calculated from monthly aggregated data, while mean elevation values represent village-level averages extracted from the 30-m Shuttle Radar Topography Mission (SRTM) digital elevation model. The long-term climatic and topographic characteristics of each sampling site are summarized in Table 1.

The study sites encompass a broad altitudinal gradient, ranging from lowland environments such as Camplong 2 Village (CP; 89 m asl.), Napu Village, East Sumba (SB; 177 m asl.), and Beaneno Village (MK; 219 m asl.) to highland environments represented by Fatumnasi Village (FT; 1,578 m asl.). Mean annual air temperature ranged from 21.8°C at the highest elevation site (FT) to 25.4-25.6°C at lowland sites (CP and MK), reflecting a strong elevation–temperature relationship. Mean annual precipitation varied from 1,452 mm at CP to 2,012 mm at Napu Village, East Sumba (SB), with higher rainfall generally observed at higher elevations and on Sumba Island. Mean annual relative humidity followed a similar elevational pattern, with higher values recorded at highland sites.

Climatic and topographic variables are presented to provide an environmental context for the sampling sites. These variables were not incorporated as explanatory factors in the morphometric analyses, and no statistical tests of trait-environment associations were performed. Accordingly, any references to environmental influences are interpreted as exploratory and descriptive rather than inferential.

All locations experience a distinct monsoonal climate, characterized by a unimodal rainfall regime with a pronounced wet season from December to March and a prolonged dry season from May to October, as illustrated in Figure 1. Seed collection was conducted with the knowledge and verbal consent of local farmers and village authorities at each site. The study did not involve protected species or restricted areas, and no formal collection permits were required under local regulations.

## Procedures

Mature arbila pods were collected from naturally growing plants at each sampling site. The species occurred predominantly in unmanaged areas, commonly near forest margins or within forested landscapes, reflecting its wild or semi-wild growth habit. Pods were harvested at full maturity, and seeds were manually extracted, air-dried under ambient conditions, and cleaned prior to analysis.

Field sampling followed a cluster-random approach based on established germplasm survey procedures (Mau et al. 2025). The primary sampling unit was defined as a location-based accession, representing a composite seed sample pooled from multiple arbila plants (approximately 3-10 individuals) within a localized area associated with a single farmer or household. These plants were typically located near farmer residences or in adjacent unmanaged areas and forest margins. As such, accession may represent a mixture of genotypes within a localized population, and all analyses were conducted at the accession level rather than at the level of individual plants.

Within each district, three to ten local arbila population centers were identified based on field reconnaissance and information provided by local farmers. From each population center, up to ten sampling locations were surveyed, according to arbila availability. Sampling locations were generally spaced approximately 500 m to 1 km apart to reduce the likelihood of sampling the same individual or closely related genets. Local farmers guided field navigation between sampling locations to ensure that each accession represented a distinct sampling unit. The number of seeds collected per accession varied depending on availability, generally ranging from 150 to 750 seeds.

Morphological characterization followed standard descriptors for grain legumes and included both qualitative and quantitative seed traits. Qualitative descriptors included seed shape, primary seed coat color, secondary seed coat color, secondary seed coat pattern, seed coat pattern, hilum shape, testa texture, and eye color, based on established descriptor guidelines with slight modifications (IBPGR 1982; Bria et al. 2019). Quantitative parameters comprised seed length (mm), seed width (mm), seed thickness (mm), and 100-seed weight (g). The full set of qualitative and quantitative seed traits evaluated in this study is presented in Table 2.

Seed length, width, and thickness were measured for each accession using technical replicates based on subsampling. Specifically, three subsamples were prepared per accession, each consisting of 50 randomly selected seeds drawn from the same composite sample. These measurements were used to estimate mean trait values and within-accession variability. The 100-seed weight was determined using one to three technical replicates per accession, depending on seed availability, and standard deviation was calculated only when two or more replicates were available. In cases of limited seed availability, measurements were conducted using the maximum possible number of complete 100-seed units.

Because subsamples were derived from the same composite accession, they were not treated as independent biological replicates. Accordingly, variation among subsamples reflects within-accession variability, whereas differences among accessions represent among-accession variation. All quantitative traits were summarized as mean values at the accession level prior to analysis. In total, 12 seed characters were evaluated. The distribution of qualitative traits is presented in Table 3. Each accession was photographed to document phenotypic characteristics and to support visual comparison among accessions.

## Data analysis

Multivariate analyses were employed as exploratory tools to summarize phenotypic structure among arbila accessions, without attempting inferential or causal interpretation. Because the sampling design does not include independent biological replication within locations, all quantitative traits were summarized at the accession level prior to analysis. Consequently, no inferential statistical tests (e.g., analysis of variance) were performed, and all results are interpreted descriptively.

A dataset comprising 108 arbila accessions collected from six locations in East Nusa Tenggara Province was compiled. The data matrix consisted of twelve seed traits, including eight qualitative characters (seed shape, primary seed coat color, secondary seed coat color, secondary seed coat pattern, seed coat pattern, hilum shape, testa texture, and eye color) and four quantitative characters (seed length, seed width, seed thickness, and 100-seed weight).

Qualitative traits were visually assessed and encoded as ordinal numerical scores based on their phenotypic categories, following established descriptor guidelines for grain legumes (IBPGR 1982; PPVTPP 2021; Mau et al. 2025). Quantitative traits were expressed as mean values per accession derived from replicated measurements. Prior to multivariate analysis, all variables were standardized using a z-score transformation to remove scale effects and ensure equal weighting among traits measured in different units. Because qualitative traits represent ordinal categories rather than continuous measurements, the use of Euclidean distance and PCA may impose artificial linear spacing among trait categories. Therefore, analyses involving qualitative variables were interpreted descriptively and limited to pattern recognition rather than as reflecting true metric relationships among categories.

To assess linear associations among quantitative seed traits, pairwise Pearson correlation coefficients were calculated across all accessions (n = 108), with significance evaluated at $\alpha = 0.05$. The correlation matrix and associated p-values were computed using PAST version 5.3 (Hammer et al. 2001).

**Table 2.** Qualitative and quantitative seed traits of arbila (*Phaseolus lunatus*) accessions

| Characteristics | Character expression and score |
|---|---|
| Seed shape | 1=round; 2=oval cuboid; 3=kidney shape |
| Seed primary color | 1=white; 2=gray; 3=yellow; 4=brown; 5=maroon; 6=purple reddish; 7=black |
| Seed secondary color | 0=no color; 1=white; 2=light brown; 3=dark brown; 4=red; 5=purple reddish |
| Secondary seed coat pattern | 0=no pattern; 1=line; 2=specks; 3=blotches; 4=mixture |
| Seed coat pattern | 0=no pattern; 1=pattern around eye only; 2=eye distinct with few specks on body; 3=eye with blotches on body; 4 = eye linked to other parts of pattern, blotches and some specks may be present; 5=eye not clean with soft specks |
| Hilum shape | 1=round; 2=ellipse;3=oval; 4=lancet |
| Testa texture | 3=smooth; 5=moderately ridged; 7=markedly ridged |
| Eye color | 1=white; 2=brown; 3=light brown |
| Seed length (mm) | measured |
| Seed width (mm) | measured |
| Seed thickness (mm) | measured |
| 100 seed weight (g) | measured |

Note: Qualitative descriptors included seed shape, primary seed coat color, secondary seed coat color, secondary pattern, seed coat pattern, hilum shape, testa texture, and eye color, following slight modifications of established descriptor guidelines (IBPGR 1982; Bria et al. 2019). In total, 12 seed characters were evaluated. Each accession was photographed to document phenotypic characteristics and support visual comparison

Principal Component Analysis (PCA) was conducted using the correlation matrix of the standardized dataset to allow variables measured on different scales to contribute equally to the ordination. PCA biplots were constructed to visualize accession scores and trait loadings across the PC1-PC2, PC1-PC3, and PC1-PC4 planes. Vector direction and length indicate the contribution of individual traits to the principal component axes, enabling simultaneous interpretation of accession distribution and the main drivers of phenotypic variation. To assess the robustness of the ordination, an additional PCA was performed using only the four quantitative seed traits (seed length, seed width, seed thickness, and 100-seed weight). This analysis was conducted on the correlation matrix of z-score standardized data to evaluate whether the primary axes of variation observed in the full dataset were consistent when the influence of ordinally coded qualitative variables was removed. Consistency between the two ordinations supports the stability of the main phenotypic gradients identified.

Hierarchical cluster analysis was conducted using Euclidean distance derived from the standardized dataset and the Unweighted Pair Group Method with Arithmetic mean (UPGMA) as the linkage algorithm. The dendrogram was truncated at a linkage distance of 15.0 to facilitate interpretation of major phenotypic patterns, particularly to distinguish highly divergent accessions from the main cluster. All statistical analyses were performed using PAST version 5.2 software (Hammer et al. 2001).

This study was designed as an exploratory analysis aimed at describing overall phenotypic variation among arbila accessions. The total number of 108 accessions was considered sufficient for this purpose, as the focus was not on statistical inference but on identifying general patterns of variation through multivariate approaches. Accordingly, the analysis emphasizes phenotypic structure and dominant axes of variation rather than hypothesis testing or estimation of population parameters.

Each accession represents a composite sample collected from multiple individual plants within a localized area and therefore does not constitute an independent biological replicate. As a result, inferential statistical analyses were not applied, and the study was limited to descriptive and multivariate approaches, including correlation analysis, Principal Component Analysis (PCA), and hierarchical clustering. The analytical workflow followed a sequential process beginning with the collection of composite accessions, followed by measurement of qualitative and quantitative seed traits and data standardization using z-score transformation. This was then followed by principal component analysis to identify dominant axes of variation and hierarchical clustering (UPGMA) to summarize phenotypic similarity among accessions. Overall, this approach was intended to detect general patterns of variation and to assess whether phenotypic diversity is structured as continuous gradients or discrete groups. To provide a clearer overview of the analytical procedure, the workflow of this study is summarized in Figure 2.

## RESULTS AND DISCUSSION

### Qualitative seed characteristics

A total of 108 arbila accessions were collected and analyzed from six locations in East Nusa Tenggara Province, representing observable phenotypic diversity within local populations (Table 3). The largest number of accessions originated from Lasi (29 accessions), followed by South Amfoang and Camplong (28 accessions each), while fewer accessions were obtained from Fatumnasi (10), Malaka (8), and East Sumba (5).

Qualitative seed traits exhibited strong polymorphism across accessions (Table 3). Seed coat coloration showed considerable diversity, with brown, gray, and black as the most frequent primary colors. Secondary coloration, particularly purple-reddish pigmentation, together with blotched and speckled patterns, was evident. These variations in seed morphology and pigmentation are illustrated in Figure 3.

Variation in seed coat pattern, hilum shape, and testa texture further contributed to phenotypic diversity. Blotched and speckled patterns were common, the hilum shape was predominantly lancet, and the testa texture was mainly smooth. Eye color exhibited limited variation, with brown dominating across accessions.

At the island level, qualitative traits showed broadly similar patterns between Timor and Sumba, with no distinct island-specific classes. Most variation occurred within rather than between islands, indicating weak geographic structuring.

Seed shape was dominated by oval cuboid and kidney-shaped forms, with oval cuboid seeds accounting for 55 accessions (50.9%) and kidney-shaped seeds for 52 accessions (48.1%), while round seeds were rare (0.9%). Seed primary color exhibited high diversity, with brown (23.1%) and gray (23.1%) being the most frequent classes, followed by black (15.7%), maroon (11.1%), white (11.1%), purple reddish (9.3%), and yellow (6.5%). Additional variation was evident in secondary seed color, with purple-reddish pigmentation recorded in 53 accessions (49.1%), whereas other colors occurred at lower frequencies. Secondary pattern expression differed among accessions, with the absence of pattern (42.6%) being most common, followed by blotches (25.0%), specks (21.3%), mixture patterns (6.5%), and line patterns (4.6%).

Further variation was present in seed coat pattern. The most frequent pattern was eye with blotches on the seed body (41.7%), followed by eye distinct with few specks on the body (34.3%). Seeds without any coat pattern accounted for 15.7%, while eye-linked patterns were recorded in 7.4% of accessions. Hilum shape was dominated by the lancet form (51.9%), followed by oval (29.6%), elliptical (16.7%), and round (1.9%) shapes. Testa texture was predominantly smooth (69.4%), with moderately ridged (29.6%) and markedly ridged (0.9%) textures occurring less frequently. Eye color exhibited limited variation, with brown eyes dominating (92.6%).

At the island level, the distribution of major qualitative seed traits showed broadly similar patterns between Timor Island (AF, CP, FT, LS, MK) and Sumba Island (SB), with no discrete island-specific trait classes observed. Oval cuboid and kidney-shaped seeds predominated on both islands, although oval cuboid forms were slightly more frequent among Timor accessions, while kidney-shaped seeds were proportionally more common in the limited Sumba sample.

Primary seed coat color classes were shared across islands, with brown, gray, and black seeds occurring in both Timor and Sumba populations. Darker pigmentation classes (brown to black) dominated on both islands, whereas lighter colors such as white and yellow were relatively rare. Secondary seed coat coloration, particularly purple-reddish pigmentation and blotched or speckled patterns, was detected at all locations, indicating that these traits are widespread rather than geographically restricted.

At the island level, comparisons suggest weak geographic structuring of qualitative seed traits, with most variation occurring within rather than between islands. These patterns are consistent with the continuous phenotypic gradients identified by PCA and cluster analysis, and support the interpretation that arbila populations in East Nusa Tenggara represent largely wild or semi-managed germplasm with limited spatial differentiation based on visible seed traits.

**Table 3.** Distribution of qualitative seed traits of arbila (*Phaseolus lunatus*) across locations in Timor and Sumba Islands, East Nusa Tenggara, Indonesia

| **Traits** | **Phenotypic class** | **Location*** | | | | | | **Total (n)** | **(%)** |
|---|---|---|---|---|---|---|---|---|---|
| | | **AF** | **CP** | **FT** | **LS** | **MK** | **SB** | | |
| Seed shape | 1=round | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0.9 |
| | 2=oval cuboid | 17 | 20 | 5 | 10 | 2 | 1 | 55 | 50.9 |
| | 3=kidney shape | 11 | 8 | 4 | 19 | 6 | 4 | 52 | 48.1 |
| Seed primary color | 1=White | 0 | 1 | 0 | 3 | 8 | 0 | 12 | 11.1 |
| | 2=gray | 4 | 14 | 0 | 4 | 0 | 3 | 25 | 23.1 |
| | 3=yellow | 3 | 0 | 2 | 2 | 0 | 0 | 7 | 6.5 |
| | 4=brown | 11 | 4 | 2 | 8 | 0 | 0 | 25 | 23.1 |
| | 5=maroon | 3 | 5 | 1 | 2 | 0 | 1 | 12 | 11.1 |
| | 6=purple reddish | 4 | 0 | 1 | 4 | 0 | 1 | 10 | 9.3 |
| | 7=black | 3 | 4 | 4 | 6 | 0 | 0 | 17 | 15.7 |
| Seed secondary color | 0=no color | 1 | 5 | 2 | 2 | 0 | 0 | 10 | 9.3 |
| | 1=white | 1 | 1 | 0 | 4 | 1 | 0 | 7 | 6.5 |
| | 2=light brown | 0 | 0 | 3 | 1 | 3 | 4 | 11 | 10.2 |
| | 3=dark brown | 4 | 4 | 3 | 5 | 0 | 0 | 16 | 14.8 |
| | 4=red | 4 | 1 | 1 | 5 | 0 | 0 | 11 | 10.2 |
| | 5=purple reddish | 18 | 17 | 1 | 12 | 4 | 1 | 53 | 49.1 |
| Secondary seed coat pattern | 0=no pattern | 11 | 15 | 3 | 12 | 3 | 2 | 46 | 42.6 |
| | 1=line | 0 | 0 | 0 | 4 | 0 | 1 | 5 | 4.6 |
| | 2=specks | 7 | 4 | 4 | 4 | 3 | 1 | 23 | 21.3 |
| | 3=blotches | 6 | 6 | 3 | 9 | 2 | 1 | 27 | 25.0 |
| | 4=mixture | 4 | 3 | 0 | 0 | 0 | 0 | 7 | 6.5 |
| Seed coat pattern | 0=no pattern | 1 | 7 | 3 | 1 | 2 | 3 | 17 | 15.7 |
| | 1=pattern around the eye only | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0.9 |
| | 2=eye distinct with few specks on the body | 7 | 10 | 5 | 9 | 5 | 1 | 37 | 34.3 |
| | 3=eye with blotches on the body | 19 | 9 | 1 | 15 | 1 | 0 | 45 | 41.7 |
| | 4=eye linked to other parts of pattern, blotches, and some specks may be present | 1 | 2 | 1 | 4 | 0 | 0 | 8 | 7.4 |
| | 5=eye not clean with soft specks | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 |
| Hilum shape | 1=round | 0 | 0 | 1 | 1 | 0 | 0 | 2 | 1.9 |
| | 2=ellipse | 6 | 7 | 4 | 1 | 0 | 0 | 18 | 16.7 |
| | 3=oval | 9 | 3 | 4 | 9 | 4 | 3 | 32 | 29.6 |
| | 4=lancet | 13 | 18 | 1 | 18 | 4 | 2 | 56 | 51.9 |
| Testa texture | 3=smooth | 25 | 26 | 5 | 15 | 3 | 1 | 75 | 69.4 |
| | 5=moderately ridged | 2 | 2 | 5 | 14 | 5 | 4 | 32 | 29.6 |
| | 7=markedly ridged | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0.9 |
| Eye color | 1=white | 0 | 1 | 0 | 1 | 0 | 0 | 2 | 1.9 |
| | 2=brown | 28 | 25 | 7 | 28 | 7 | 5 | 100 | 92.6 |
| | 3=light brown | 0 | 2 | 3 | 0 | 1 | 0 | 6 | 5.6 |

Note: Location*: AF: Amfoang, CP: Camplong 2, FT: Fatumnasi, LS: Lasi, MK: Malaka, SB: East Sumba

**Quantitative seed characteristics**

Quantitative seed traits showed wide variation among the 108 arbila accessions (Table 4). Seed length ranged from 10.33 to 23.70 mm, seed width from 6.74 to 19.71 mm, and seed thickness from 2.34 to 13.98 mm. The 100-seed weight varied from 25.40 to 56.50 g, reflecting marked differences in seed size and mass across accessions.

Pearson correlation analysis revealed significant positive associations among all pairs of quantitative seed traits (Table 5; $p < 0.0001$ for all pairs). Seed length and seed width were very strongly correlated ($r = 0.94$), and seed width likewise showed a strong association with seed thickness ($r = 0.90$). The association between seed length and seed thickness was also substantial ($r = 0.79$), indicating that the three linear dimensions co-vary closely across accessions. By contrast, 100-seed weight exhibited comparatively weaker correlations with seed length ($r = 0.48$), seed width ($r = 0.41$), and seed thickness ($r = 0.25$), suggesting that factors beyond external linear dimensions, such as seed density or degree of seed fill, influence seed mass.

These strong intercorrelations among dimensional traits are consistent with their joint contribution to PC2, confirming that this axis primarily represents variation in linear seed dimensions. In contrast, PC1 is dominated by 100-seed weight, indicating that seed mass constitutes the primary axis of phenotypic differentiation. The relatively weaker association of 100-seed weight with the dimensional traits shows that seed mass constitutes a partially independent morphological component, which contributes to the variation captured by secondary principal components.

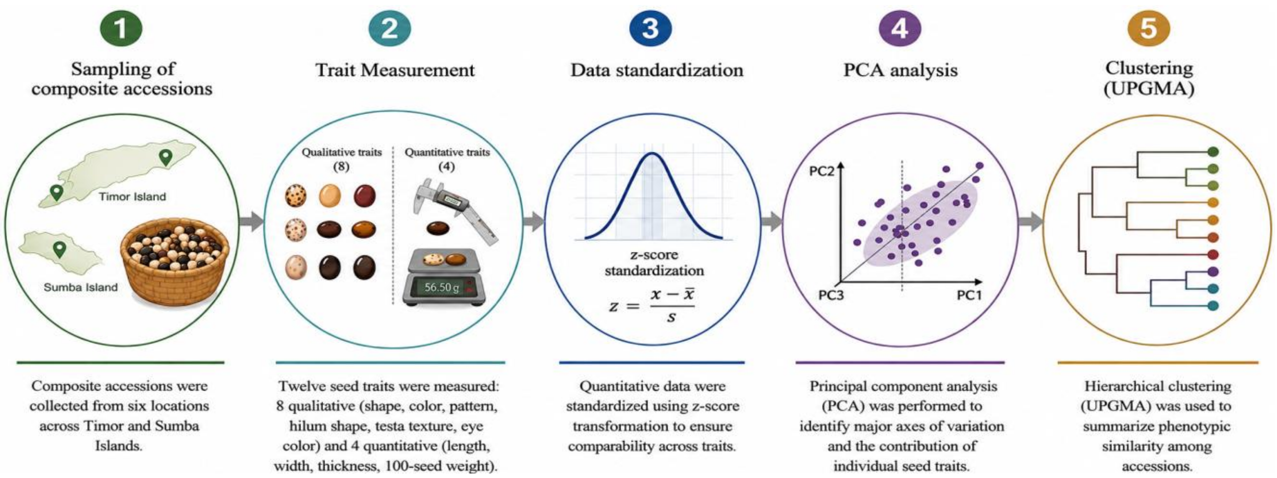


**Figure 2.** Analytical workflow of the study showing the sequential steps from composite accession sampling, trait measurement, data standardization, Principal Component Analysis (PCA), and hierarchical clustering (UPGMA)

**Figure 3.** Representative seed samples of poisonous arbila (*Phaseolus lunatus*) collected from different locations in East Nusa Tenggara, Indonesia

**Table 4.** Quantitative seed characteristics of arbila (*Phaseolus lunatus*) accessions, showing mean (± SD) by sampling location

| Location | n | Seed length (mm) | Seed width (mm) | Seed thickness (mm) | 100-seed weight (g) |
|---|---|---|---|---|---|
| Amfoang (AF) | 28 | 13.16 ± 0.84 | 8.45 ± 0.77 | 3.57 ± 0.73 | 37.04 ± 1.21 |
| Camplong (CP) | 28 | 14.68 ± 3.04 | 10.81 ± 3.03 | 5.82 ± 2.79 | 37.11 ± 1.04 |
| Fatumnasi (FT) | 10 | 13.09 ± 0.36 | 9.25 ± 0.21 | 4.85 ± 0.58 | 38.97 ± 1.66 |
| Lasi (LS) | 29 | 12.79 ± 1.01 | 8.95 ± 0.57 | 4.60 ± 0.52 | 38.06 ± 1.41 |
| Malaka (MK) | 8 | 11.78 ± 0.55 | 7.76 ± 0.49 | 4.06 ± 0.49 | 33.73 ± 1.82 |
| East Sumba (SB) | 5 | 12.67 ± 0.44 | 8.48 ± 0.15 | 4.37 ± 0.32 | 32.31 ± 1.13 |
| Total/all accessions | 108 | 12.87 ± 3.14 | 8.92 ± 2.59 | 4.53 ± 1.95 | 35.93 ± 1.64 |

**Table 5.** Pearson correlation coefficients among quantitative seed traits of arbila (*Phaseolus lunatus*) accessions (n = 108)

| Trait | Seed length | Seed width | Seed thickness | 100-seed weight |
|---|---|---|---|---|
| Seed length | 1 | 0.94** | 0.79** | 0.48** |
| Seed width | 0.94** | 1 | 0.90** | 0.41** |
| Seed thickness | 0.79** | 0.90** | 1 | 0.25** |
| 100-seed weight | 0.48** | 0.41** | 0.25** | 1 |

**Table 6.** Variance explained by the first twelve principal components derived from PCA of twelve seed traits of arbila (*Phaseolus lunatus*) accessions

| PC | % variance | PC | % variance |
|---|---|---|---|
| 1 | 65.2 | 7 | 1.3 |
| 2 | 13.8 | 8 | 0.9 |
| 3 | 7.8 | 9 | 0.8 |
| 4 | 4.8 | 10 | 0.4 |
| 5 | 3.1 | 11 | 0.2 |
| 6 | 1.6 | 12 | 0.1 |

**Table 7**. Trait loadings for the first four principal components (PC1-PC4) based on PCA of twelve seed traits of arbila (*Phaseolus lunatus*) accessions

| Trait | PC1 | PC2 | PC3 | PC4 |
|---|---|---|---|---|
| Seed shape | -0.009 | 0.004 | 0.046 | -0.004 |
| Seed primary color | 0.013 | -0.197 | **-0.775** | **0.526** |
| Seed secondary color | -0.080 | -0.137 | 0.457 | **0.671** |
| Secondary seed coat pattern | -0.014 | -0.095 | 0.376 | 0.109 |
| Seed coat pattern | -0.008 | -0.146 | 0.189 | 0.362 |
| Hilum shape | 0.004 | -0.018 | 0.076 | 0.059 |
| Testa texture | -0.013 | -0.026 | 0.040 | -0.199 |
| Eye color | 0.000 | -0.009 | -0.011 | 0.027 |
| Seed length | 0.178 | **0.504** | 0.007 | 0.219 |
| Seed width | 0.158 | **0.560** | -0.023 | 0.184 |
| Seed thickness | 0.098 | **0.530** | 0.000 | 0.064 |
| 100-seed weight | **0.963** | -0.251 | 0.059 | -0.027 |

Note: Only PC1-PC4 are shown, as these components collectively explained 91.6% of the total phenotypic variance, whereas subsequent components contributed marginally and were not interpreted. Loadings with absolute values ≥ 0.50 are highlighted in bold to indicate major contributions, regardless of sign

## Principal Component Analysis (PCA)

Principal Component Analysis (PCA) of the twelve seed traits yielded twelve components, of which the first four together explained 91.6% of the total standardized variance (Table 6). PC1 accounted for 65.2% of the variance and was dominated by 100-seed weight (loading = 0.963; Table 7), indicating that seed mass represents the primary axis of phenotypic differentiation among arbila accessions. Seed length and seed width also contributed positively, but to a much lesser extent.

PC2 explained an additional 13.8% of the variance and was primarily defined by the three linear dimensional traits: seed length (0.504), seed width (0.560), and seed thickness (0.530; Table 7). This component captures variation in seed morphology that is largely independent of seed mass. Together, PC1 and PC2 accounted for 79.0% of the total variance and represent the dominant axes of morphometric variation.

PC3 (7.8%) and PC4 (4.8%) were mainly associated with pigmentation-related traits (Table 7). PC3 was largely influenced by primary seed coat color, with additional contributions from secondary color and pattern, whereas PC4 was defined mainly by variation in seed coat coloration. Because these traits were recorded as ordinal categories, their contributions are interpreted descriptively as reflecting differences among phenotypic classes rather than precise quantitative distances.

The PCA biplots (Figures 4.A-C) illustrate the distribution of accessions and the relative contributions of seed traits across the PC1-PC2, PC1-PC3, and PC1-PC4 planes. Together, these ordinations facilitate visualization of the primary morphometric gradients, with PC1 largely reflecting variation in seed mass and PC2 representing differences in linear seed dimensions, while the secondary variation associated with pigmentation-related traits is captured by PC3 and PC4. Seed weight is closely aligned with PC1, whereas seed length, width, and thickness are associated with PC2, indicating that seed mass and linear dimensions are not entirely coupled. Qualitative traits are positioned near the origin, suggesting a relatively minor influence on the primary axes of variation.

Accessions are distributed along a phenotypic gradient rather than forming distinct groups. Those from Camplong and Lasi span a broad range along PC1, whereas accessions from Malaka and East Sumba are mostly located toward the lower end of this axis. Along PC2, accessions are widely dispersed without clear separation among locations, indicating overlapping ranges of seed dimensions across sites. Overall, these results suggest that phenotypic variation in arbila is primarily structured by seed size-related traits, forming a dominant size-related gradient, while qualitative traits contribute mainly to secondary variation.

### Exploratory ecological visualization (CCA)

This analysis is presented as an exploratory visualization of potential associations and does not imply causal or statistically tested relationships. Canonical Correspondence Analysis (CCA) was used as an exploratory ordination to visualize potential associations between seed traits and environmental descriptors without testing statistical significance or inferring causal relationships (Figure 5). This visualization is intended to provide ecological context rather than formal inference. Unlike PCA, which summarizes variation within a single dataset, CCA provides a visual representation of potential associations between trait variables and environmental factors. The resulting biplot illustrates the relative orientation of traits and environmental variables, where vector direction and proximity indicate potential associations. The resulting biplot illustrates the relative orientation of traits and environmental variables, where vector direction and proximity indicate potential associations. Accordingly, these relationships should be interpreted descriptively rather than inferentially.

### Hierarchical cluster analysis

Hierarchical cluster analysis based on Euclidean distance and the UPGMA method was used to summarize phenotypic similarity among arbila accessions. When the dendrogram was truncated at a distance of 15.0, corresponding to a clear increase in fusion distance, three clusters were identified (Figure 6 and Table 8), although their sizes were highly uneven.

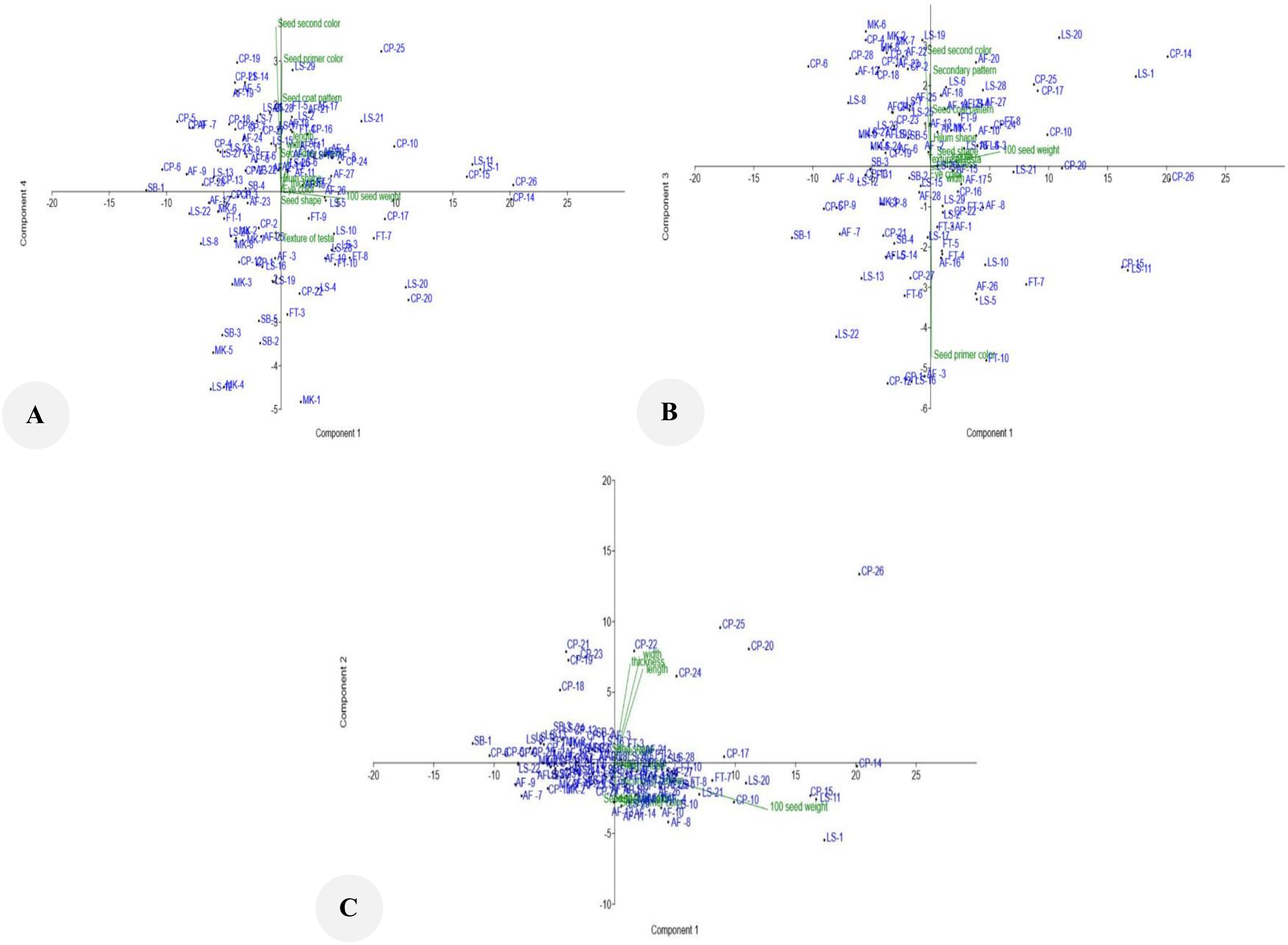


**Figure 4.** Principal Component Analysis (PCA) biplots showing the distribution of 108 arbila (*Phaseolus lunatus*) accessions and the relative positions of seed trait loadings based on twelve seed traits. The ordinations are presented as: A. PC1 versus PC2, B. PC1 versus PC3, and C. PC1 versus PC4. Points represent individual accessions identified by accession codes corresponding to their sampling locations (AF, CP, FT, LS, MK, and SB). Green trait labels indicate the relative contributions of seed characters to the principal component axes

**Table 8.** Distribution of arbila (*Phaseolus lunatus*) accessions across clusters defined by hierarchical cluster analysis at a truncation distance of 15.0

| Cluster | Accession codes | Number of accessions |
|---|---|---|
| Cluster I | CP-26 | 1 |
| Cluster II | CP-14, CP-15, LS-1, LS-11 | 4 |
| Cluster III | AF-1 to AF-28, CP-1 to CP-13, CP-16 to CP-25, CP-27, CP-28, LS-2 to LS-10, LS-12 to LS-29, FT-1 to FT-10, MK-1 to MK-8, SB-1 to SB-5 | 103 |

Note: Cluster sizes are highly uneven, reflecting a continuous phenotypic gradient rather than discrete group separation

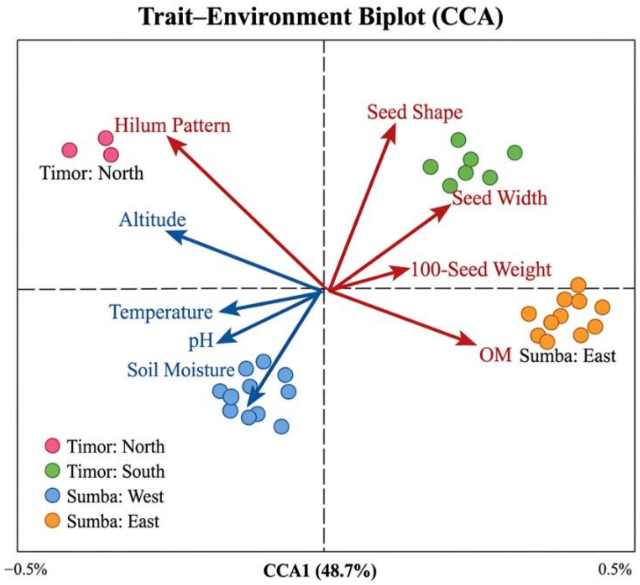


**Figure 5.** Exploratory trait-environment ordination based on Canonical Correspondence Analysis (CCA), illustrating the relative orientation of seed morphometric traits (red vectors) and environmental variables (blue vectors) across arbila accessions. Vector direction and proximity indicate potential associations and should be interpreted descriptively rather than as statistically tested relationships

Cluster I consisted of a single highly divergent accession (CP-26), which separated from all other accessions at a relatively large linkage distance, reflecting marked phenotypic distinctness. Cluster II included four accessions (CP-14, CP-15, LS-1, and LS-11) that diverged earlier than most accessions but remained closer to the main group than CP-26. In contrast, Cluster III comprised the remaining 103 accessions, which grouped at shorter linkage distances and formed a large, relatively homogeneous cluster.

Overall, this pattern suggests that phenotypic variation among arbila accessions is largely continuous, with only a small number of clearly divergent individuals. This interpretation is consistent with the PCA results (Figure 4), which suggest a dominant gradient driven largely by seed size-related traits. Because the analysis includes both quantitative traits and ordinally encoded qualitative variables, the results are interpreted descriptively. The general structure remained stable across nearby truncation levels, supporting the robustness of the clustering pattern.

## Discussion

The present study documents substantial phenotypic diversity among 108 poisonous arbila accessions collected from six locations in East Nusa Tenggara Province. Both qualitative and quantitative seed traits varied widely, reflecting the predominantly wild or semi-managed status of arbila populations across heterogeneous environments. Comparable levels of morphological variation have been reported for lima bean landraces in Brazil, Benin, Côte d'Ivoire, and Timor Island, supporting the use of seed-based traits as reliable descriptors of intraspecific diversity in *P. lunatus* under contrasting ecological conditions (Bria et al. 2019; Sandrine et al. 2020; Fagbédji et al. 2023; Barros et al. 2024).

The predominance of oval, cuboid, and kidney-shaped seeds, together with the wide range of seed size observed in this study, is consistent with patterns reported from heterogeneous dryland environments characterized by irregular rainfall and low soil fertility. In marginal environments, seed size and shape are often linked to germination and early seedling establishment, as larger seeds can provide greater reserves under resource-limited conditions. However, in this study, such interpretations are based solely on general patterns reported in the literature and cannot be considered as direct evidence of adaptation, as no formal statistical tests of trait-environment associations were conducted (Sita et al. 2017; Dave et al. 2024).

In addition, seed coat coloration contributed substantially to the overall phenotypic diversity. The predominance of darker primary seed coat colors, particularly brown, gray, and black, is consistent with findings in other dryland legumes, where darker pigmentation has been associated with stress tolerance and farmer preferences (Bria et al. 2019; Mau et al. 2025). Previous studies have linked variation in seed coat color to differences in biochemical composition, including flavonoids and anthocyanins, which may influence ecological interactions and stress responses (Tudor et al. 2024; Plestenjak et al. 2025). In this study, however, such associations are discussed only as potential explanations and are not directly inferred from the data.

Seed coat pigmentation and pattern traits exhibited pronounced polymorphism. The frequent occurrence of purple-reddish secondary coloration combined with specks and blotches suggests complex genetic control, potentially involving multiple interacting loci. Similar phenotypic variation in other legumes has been associated with differences in flavonoid and proanthocyanidin pathways, as well as functional roles related to tolerance to ultraviolet radiation, oxidative stress, and microbial interactions.

However, in the current analysis, these genetic and functional interpretations are considered hypothetical and are based on previous research rather than direct measurements of arbila accessions (Quilichini et al. 2022; Islam and Dhaubhadel 2023; Jaiswal and Dakora 2024).

Comparative studies across grain legumes imply that pigmentation traits often show phenotypic variation in wild and semi-managed populations. Environmental heterogeneity, combined with potential gene exchange among populations, may prevent the fixation of discrete phenotypic classes. The patterns found here appear to be consistent with other *Phaseolus* systems, where phenotypic structure reflects a combination of environmental influences and shared evolutionary history rather than sharply defined groups (Smýkal et al. 2020; Wang et al. 2023).

From an evolutionary perspective, variation in seed size and pigmentation has often been linked to early domestication processes in legumes (Ku et al. 2020; Chen et al. 2021). The presence of relatively large seeds and highly polymorphic seed coat patterns suggests that the studied arbila populations may not represent strictly wild forms. This interpretation is further supported by the ecological context, where arbila commonly occurs near both unmanaged habitats and farmer residences. Under such conditions, gene exchange among wild, semi-managed, and potentially cultivated populations is likely. Consequently, morphometric diversity likely reflects both natural variation and anthropogenic influences, including introgression and escape from cultivation (Heredia-Pech et al. 2022; Penha et al. 2025).

The multivariate analyses presented here should be interpreted as exploratory summaries of overall phenotypic structure. Because the dataset integrates quantitative traits and ordinal-coded qualitative traits, the PCA should not be interpreted as implying linear relationships or metric distances among qualitative categories, but rather as a visualization of dominant patterns of variation. The first four principal components accounted for 91.6% of the total variation (Table 6), with PC1 dominated by quantitative seed size traits, while qualitative traits contributed primarily to secondary axes.

Hierarchical cluster analysis revealed a markedly unbalanced structure, dominated by one large cluster and a small number of highly divergent accessions. Notably, accessions such as CP-26, CP-14, and LS-1 occupied extreme positions along the primary phenotypic gradient defined by PC1, corresponding to the upper range of seed size variation, whereas accessions such as CP-13 were positioned toward the lower end. This pattern is more appropriately interpreted as a continuous gradient with a limited number of outliers, rather than as discrete phenotypic groups (Fagbédji et al. 2023; Barros et al. 2024). The consistency between ordination (PCA) and clustering outcomes strengthens the view that arbila diversity is structured along a single dominant axis, rather than partitioned into clearly separated groups. These extreme phenotypes contribute substantially to the overall variance structure and may represent important targets for germplasm conservation and future selection. Accordingly, cluster delineation in this study should be regarded as heuristic and descriptive, rather than as evidence of distinct biological groupings, and the selected truncation distance does not correspond to a biologically meaningful threshold. Phenotypic variation, therefore, follows a structured and interpretable pattern rather than a random distribution, and is best understood within a broader ecological and evolutionary framework.

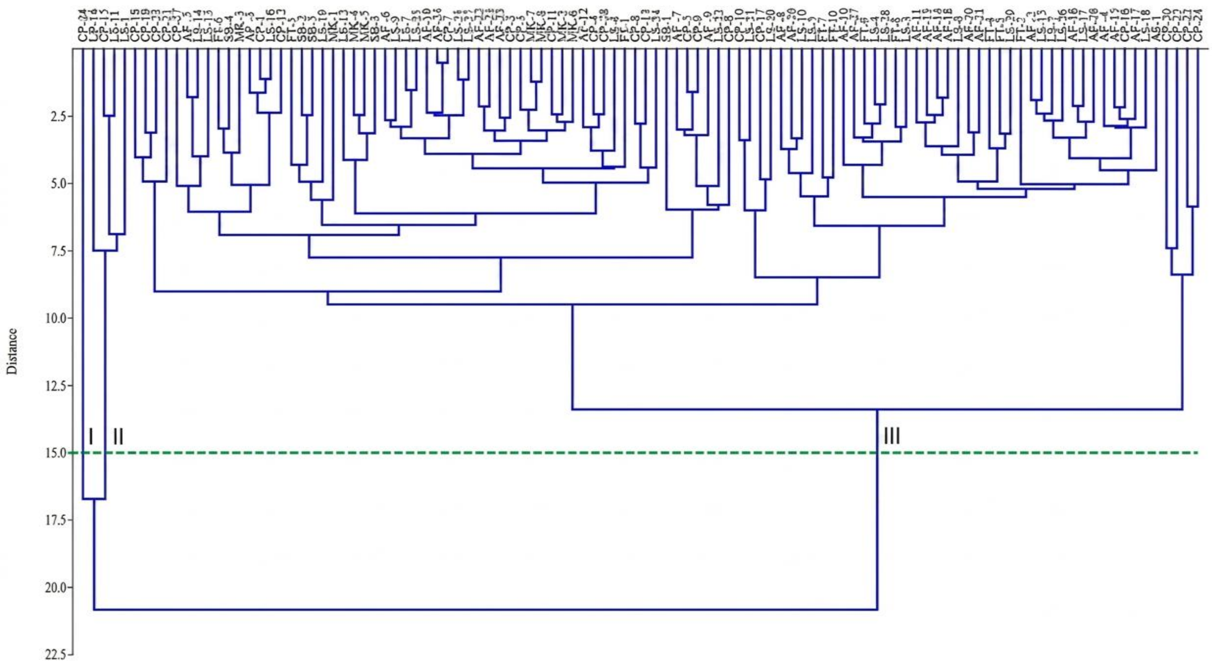


**Figure 6.** Hierarchical cluster dendrogram of 108 arbila (*Phaseolus lunatus*) accessions constructed using Euclidean distance and the UPGMA method based on twelve seed traits. The dashed horizontal line indicates a heuristic truncation distance (15.0) corresponding to a marked increase in fusion distance, highlighting the separation of a small number of highly divergent accessions from a large, dense assemblage of phenotypically similar accessions

The phenotypic patterns identified here are best interpreted within a domestication continuum framework, where arbila populations do not form a strict dichotomy between wild and cultivated types. Instead, variation in seed size, shape, and pigmentation occurring along a gradient, together with the absence of discrete clustering, suggests that arbila exists along a domestication continuum ranging from unmanaged wild populations to semi-managed forms influenced by low-intensity human selection. In the dryland agroecosystems of East Nusa Tenggara, where arbila occurs both near settlements and in natural habitats, low-intensity and non-systematic selection likely maintains phenotypic diversity rather than promoting uniform domestication traits. This results in a dynamic system shaped by environmental filtering, gene flow, and phenotypic plasticity. To place these patterns within a broader ecological and evolutionary context, and to integrate morphological, multivariate, and ecological perspectives into a unified framework, a conceptual model of the domestication continuum is presented in Figure 7.

Building on this conceptual framework, Figure 8 provides an integrated analytical synthesis linking trait variation, principal component structure, and clustering outcomes into a unified framework. The results demonstrate that seed size–related traits dominate the primary axis of variation (PC1 = 65.2%), forming a size-driven gradient that underlies the major cluster (Cluster III; 103 accessions). In contrast, qualitative traits contribute to secondary axes (PC3-PC4) and are broadly distributed without forming discrete phenotypic groups. This synthesis reinforces the interpretation that arbila diversity is structured along a continuous phenotypic gradient rather than discrete clusters.

Each accession in this study represents a composite seed sample collected from multiple individual plants within a localized area. As a result, an accession may contain a mixture of genotypes, and the observed phenotypic variation reflects both within- and among-plant variation rather than independent biological replication at the individual level. Consequently, the present design does not allow separation of genetic and environmental sources of variation, and variation within and among individual plants is inherently confounded within each accession. Therefore, inference at the individual-plant level or on true population structure is not supported. Interpretations related to population differentiation, domestication status, or selection processes should thus be regarded as descriptive and hypothesis-generating rather than conclusive. Accordingly, location-level summaries reflect variation among accessions rather than replicated population-level estimates.

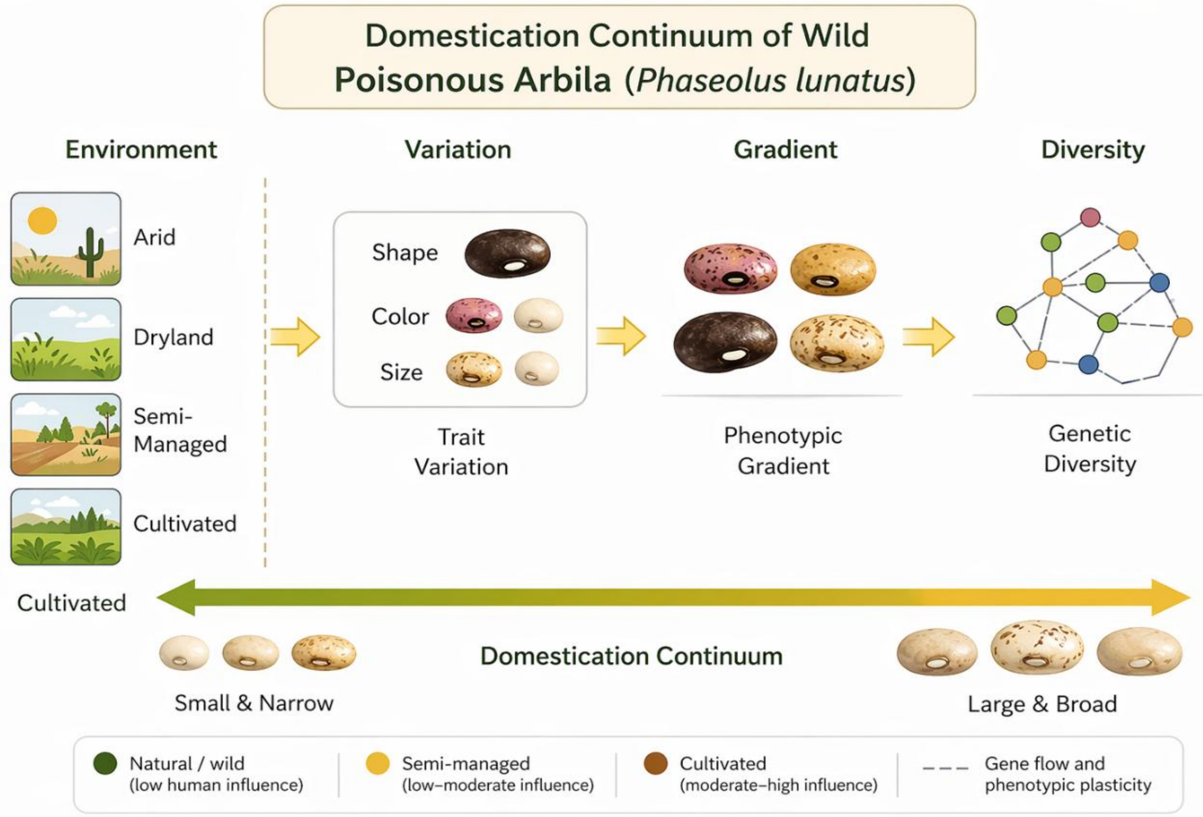
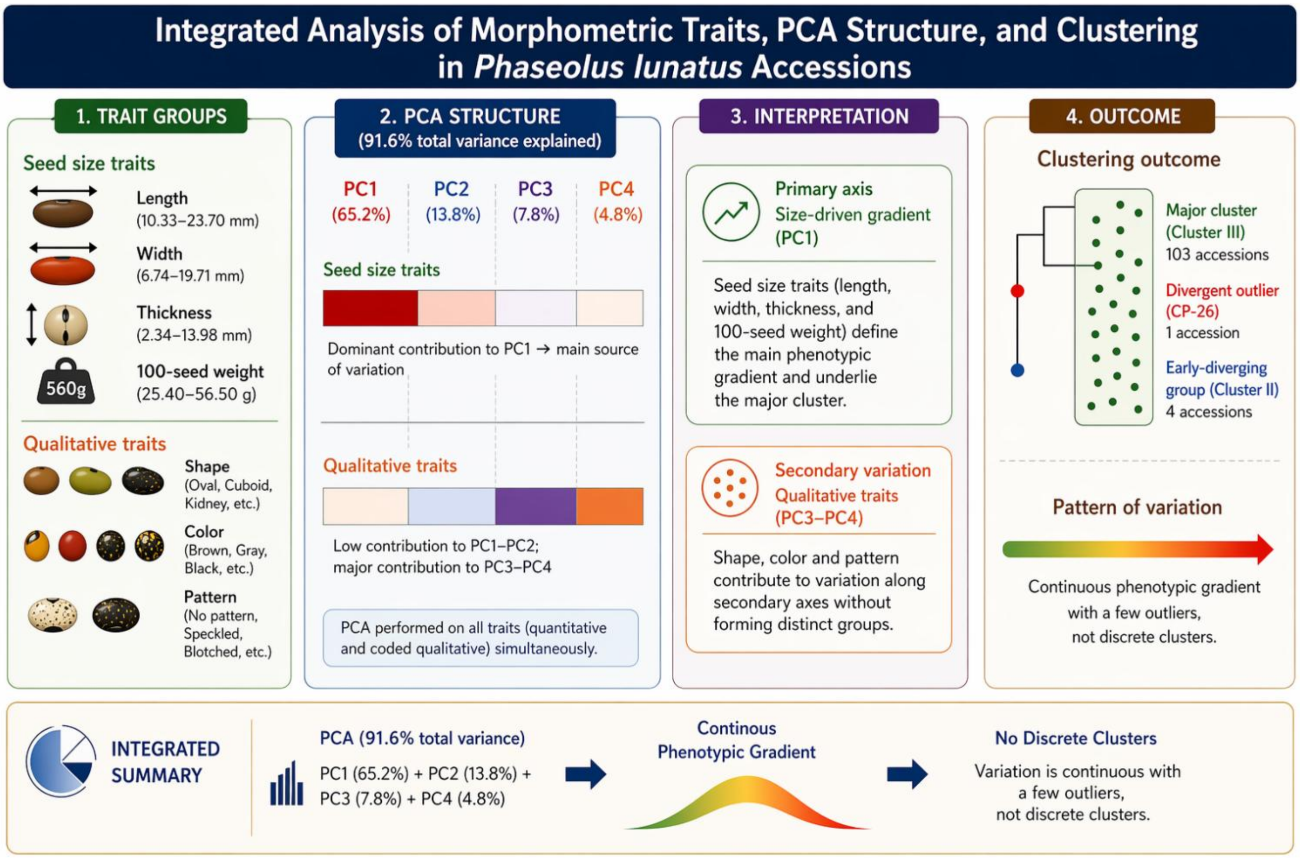


**Figure 7.** Conceptual model illustrating the domestication continuum of arbila (*Phaseolus lunatus*) across environmental and management gradients. The model links environmental conditions (arid, dryland, semi-managed, and cultivated) with trait variation (seed shape, color, and size), resulting in a phenotypic gradient and associated diversity patterns. The bidirectional arrow represents variation in seed size from small and narrow to large and broad, reflecting a continuous phenotypic gradient rather than discrete differentiation among arbila accessions. This continuum is shaped by the combined influence of environmental filtering, gene flow, and low-intensity human selection



Dominant contribution to PC1 → main source of variation

Low contribution to PC1–PC2; major contribution to PC3–PC4

PCA performed on all traits (quantitative and coded qualitative) simultaneously.

Seed size traits (length, width, thickness, and 100-seed weight) define the main phenotypic gradient and underlie the major cluster.

Shape, color and pattern contribute to variation along secondary axes without forming distinct groups.

Continuous phenotypic gradient with a few outliers, not discrete clusters.

Variation is continuous with a few outliers, not discrete clusters.

**Figure 8.** Integrated synthesis of seed trait variation, principal component structure, and clustering patterns in arbila (*Phaseolus lunatus*). Seed size-related traits dominate the primary axis (PC1 = 65.2%), forming a size-driven gradient that underlies the main cluster (Cluster III; 103 accessions). In contrast, qualitative traits contribute to secondary axes (PC3 = 7.8% and PC4 = 4.8%) and are broadly distributed without forming distinct groups

The partial correspondence between cluster structure and geographic origin suggests that arbila diversity is shaped by both spatial and non-spatial factors. More broadly, the ecological context suggests that morphometric variation does not arise solely from differentiation within isolated wild populations.

Arbila plants were frequently found near both farmer residences and unmanaged habitats, suggesting that opportunities for gene exchange among populations are likely. Under such conditions, phenotypic variation may reflect introgression among wild, semi-managed, and potentially cultivated gene pools, likely contributing to the overall diversity recorded. However, as this study is based exclusively on morphological data derived from composite accessions, the extent of gene flow cannot be determined and should be regarded as a hypothesis requiring molecular validation.

Despite the wide altitudinal range of sampling sites, no consistent relationship between seed size and elevation was detected. Although the highest mean seed weight occurred in a highland population, relatively large seeds were also found in lowland sites, indicating that seed size is not determined by elevation alone. Patterns analogous to Bergmann's rule (Atkinson 1994; Mousseau 1997) are not supported in this dataset and require further investigation. In legumes, larger seed size is often associated with improved seedling establishment under stress conditions, but this relationship was not consistently evident in the present study (Sita et al. 2017; Dave et al. 2024). While environmental variables were not formally included in the statistical models, the phenotypic gradients identified here are consistent with environmentally structured variation across heterogeneous dryland systems, suggesting a potential role of environmental filtering that warrants further investigation using inferential approaches.

Morphological data alone are insufficient to confirm gene flow, population genetic structure, or biochemical composition. To address these limitations, future studies integrating molecular and biochemical approaches will be essential. Accordingly, interpretations related to domestication status, genetic structure, and detoxification potential should be considered as working hypotheses. Robust validation will require molecular markers, such as SSRs or genome-wide SNPs, to assess genetic structure and connectivity among populations, as demonstrated in recent studies on *P. lunatus*, which revealed extensive allelic diversity and weak population structuring in wild and semi-managed germplasm (Bohra et al. 2022; Zhang et al. 2022; Paipilla et al. 2025). In addition, quantitative analyses of cyanogenic compounds and phenolic profiles are needed to directly link seed pigmentation with toxicity and nutritional attributes (Ambika et al. 2022; Zhang 2024).

Qualitative seed traits in this study were encoded as ordinal categories for inclusion in multivariate analyses. While this approach is commonly used in exploratory morphometric studies, ordinal coding does not represent true metric distances among trait categories. Consequently, the contribution of qualitative traits to PCA and clustering should be interpreted descriptively. Future studies employing mixed-data approaches, such as Gower distance, Factor Analysis of Mixed Data (FAMD), or Multiple Correspondence Analysis (MCA), would provide a more rigorous framework when inferential modeling is required.

Although the present study did not include clinical or toxicological assessments, local observations imply that improper consumption of arbila seeds may lead to acute symptoms such as vomiting and respiratory distress. These effects are consistent with cyanogenic toxicity rather than chronic metal-related poisoning. Further studies integrating toxicological, clinical, and biochemical analyses would be valuable to better understand the health implications of arbila consumption under traditional processing practices. Future studies should also consider standardized detoxification protocols and safe consumption thresholds to support the sustainable use of arbila as a food resource.

A balanced approach that considers both the use of arbila and the promotion of widely cultivated staple crops may strengthen food system resilience while minimizing potential health risks. In addition, community-based education on safe preparation methods and the cultural context of arbila consumption may help reduce stigma and support its informed use as a complementary food resource.

The number of accessions varied among locations, resulting in an uneven sampling distribution across sites. Relatively large numbers of accessions represented some locations, whereas others, particularly Sumba, were represented by fewer samples. This imbalance may influence the resulting patterns of phenotypic variation and limit the representativeness of location-level comparisons. Accordingly, differences among locations should be interpreted with caution, and the patterns described here are best considered indicative rather than definitive. More balanced sampling across sites would be necessary to support more robust geographic comparisons in future studies.

In conclusion, this study demonstrates considerable morphometric diversity among 108 poisonous arbila accessions from Timor and Sumba Islands, with variation structured along a phenotypic gradient rather than discrete groups. Multivariate analysis reveals that seed size–related traits dominate the primary axis of variation (PC1 = 65.2%), with the first four components explaining 91.6% of the total variance, whereas qualitative traits mainly contribute to secondary axes without forming distinct clusters. Hierarchical clustering supports this pattern, with most accessions grouped into a single large cluster and only a few clearly divergent outliers. These results provide a useful baseline for germplasm conservation and highlight the potential of divergent accessions for selection and early-stage domestication in dryland systems. However, the findings should be interpreted with caution due to the exploratory nature of the analysis, including the use of composite accessions, ordinal trait encoding, and the absence of molecular and inferential data. Future studies integrating molecular markers, trait–environment analyses, and biochemical characterization will be essential to validate genetic structure, assess adaptive variation, and support breeding and domestication strategies for this underutilized legume. This study contributes to the understanding of underutilized legumes in dryland agroecosystems by demonstrating that phenotypic diversity in arbila is structured along a continuous gradient shaped by interacting ecological and evolutionary processes, highlighting the importance of integrating morphological, ecological, and genetic approaches in future research.

## ACKNOWLEDGEMENTS

The authors acknowledge that this study was partially supported by the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia, through the Fundamental Research Scheme 2025 (Main Contract No. 092/C3/DT.05.00/PL/2025; Derivative Contract No. 131/UN15.22/PL/2025).

## REFERENCES

Ambika, Aski MS, Gayacharan, Hamwieh A, Talukdar A, Gupta SK, Sharma BB, Joshi R, Upadhyaya HD, Singh K, Kumar R. 2022. Unraveling origin, history, genetics, and strategies for accelerated domestication and diversification of food legumes. Front Genet 13: 932430. https://doi.org/10.3389/fgene.2022.932430.

Antriyandarti E, Barokah U, Rahayu W, Asami A, Laia DH, Sari LD, Pranadita NE, Melati NSK. 2024. Resilience of dryland farm households in the mountains and their adaptability to environmental and social challenges. Environ Chall 17: 101037. https://doi.org/10.1016/j.envc.2024.101037.

Atkinson D. 1994. Temperature and organism size: A biological law for ectotherms?. Adv Ecol Res 25: 1-58. https://doi.org/10.1016/S0065-2504(08)60212-3.

Barros ES, Sarigu M, Lallai A, Nicolau JPB, Benedito CP, Bacchetta G, Torres SB. 2024. Phenotypic identification of landraces of *Phaseolus lunatus* L. from the Northeastern Region of Brazil using morpho-colorimetric analysis of seeds. Horticulturae 10 (9): 948. https://doi.org/10.3390/horticulturae10090948.

Bohra A, Tiwari A, Kaur P, Ganie SA, Raza A, Roorkiwal M, Mir RR, Fernie AR, Smýkal P, Varshney RK. 2022. The key to the future lies in the past: Insights from grain legume domestication and improvement should inform future breeding strategies. Plant Cell Physiol 63 (11): 1554-1572. https://doi.org/10.1093/pcp/pcac086.

Bria EJ, Suharyanto E, Purnomo. 2019. Variability and intra-specific classification of lima bean (*Phaseolus lunatus* L.) from Timor Island based on morphological characters. J Trop Biodivers Biotechnol 4 (2): 62-71. https://doi.org/10.22146/jtbb.42547.

Chacón-Sánchez MI, Martínez-Castillo J. 2017. Testing domestication scenarios of lima bean (*Phaseolus lunatus* L.) in Mesoamerica: Insights from genome-wide genetic markers. Front Plant Sci 8: 1551. https://doi.org/10.3389/fpls.2017.01551.

Chen X, Zhou G, Pang J, Srinives P. 2021. Editorial: Domestication of agronomic traits in legume crops. Front Genet 12: 707600. https://doi.org/10.3389/fgene.2021.707600.

Dave K, Kumar A, Dave N, Jain M, Dhanda PS, Yadav A, Kaushik P. 2024. Climate change impacts on legume physiology and ecosystem dynamics: A multifaceted perspective. Sustainability 16 (14): 6026. https://doi.org/10.3390/su16146026.

Fagbédji RF, Djedatin LG, Nanoukon C, Kinhoegbe G, Havivi A, Yédomonhan H, Agbangla C. 2023. Insight into the genetic diversity of cultivated lima bean (*Phaseolus lunatus* L.) in Benin. Am J Mol Biol 13 (1): 32-45. https://doi.org/10.4236/ajmb.2023.131003.

Hammer Ø, Harper DAT, Ryan PD. 2001. PAST: Paleontological statistics software package for education and data analysis. Palaeontol Electron 4 (1): 1-9. https://palaeo-electronica.org.

Heredia-Pech M, Chávez-Pesqueira M, Ortiz-García MM, Andueza-Noh RH, Chacón-Sánchez MI, Martínez-Castillo J. 2022. Consequences of introgression and gene flow on the genetic structure and diversity of lima bean (*Phaseolus lunatus* L.) in its Mesoamerican diversity area. PeerJ 10: e13690. https://doi.org/10.7717/peerj.13690.

International Board for Plant Genetic Resources (IBPGR). 1982. Lima Bean Descriptors. IBPGR Secretariat, Rome.

Islam NS, Dhaubhadel S. 2023. Proanthocyanidin biosynthesis and postharvest seed coat darkening in pinto bean. Phytochem Rev 24 (4): 2445-2461. https://doi.org/10.1007/s11101-023-09895-8.

Jaiswal SK, Dakora FD. 2024. Seed-coat pigmentation plays a crucial role in partner selection and $N_2$ fixation in legume-root-microbe associations in African soils. Plants 13 (11): 1464. https://doi.org/10.3390/plants13111464.

Ku YS, Contador CA, Ng MS, Yu J, Chung G, Lam HM. 2020. The effects of domestication on secondary metabolite composition in legumes. Front Genet 11: 581357. https://doi.org/10.3389/fgene.2020.581357.

Mau YS, Oematan SS, Arsa IGBA, Ndiwa ASS, Bere AC. 2025. Genetic diversity of local cowpea (*Vigna unguiculata*) germplasms of East Nusa Tenggara Province, Indonesia, revealed by seed characteristics. Biodiversitas 26 (3): 1279-1288. https://doi.org/10.13057/biodiv/d260327.

Mousseau TA. 1997. Ectotherms follow the converse of Bergmann's rule. Evolution 51 (2): 630-632. https://doi.org/10.2307/2411138.

Naisali H, Witoyo JE, Utoro PAR. 2023. Local legumes from the dry land of East Nusa Tenggara: Diversity, nutritional composition, and their use in society: A literature study. Jurnal Ilmiah Teknologi dan Industri Pangan UNISRI 8 (2): 155-166. https://doi.org/10.33061/jitipari.v8i2.9054. [Indonesian]

Paipilla G, Escobar MC, Ovalle MF, Vanegas BS, García T, Chacón-Sánchez MI. 2025. Genome-wide discovery of short sequence repeat markers useful for genetic diversity studies in lima bean (*Phaseolus lunatus* L.). Sci Rep 16 (1): 1419. https://doi.org/10.1038/s41598-025-31509-w.

Penha JS, Oliveira GCX, Lopes ACA, Gomes RLF, Costa MF, Pinheiro JB, Zucchi MI, Assunção-Filho JR, Silvestre EA, Dequigiovanni G, Martínez-Castillo J. 2025. Genetic diversity and population structure of lima bean (*Phaseolus lunatus* L.) based on cpSSR markers. Braz J Biol 85: e298660. https://doi.org/10.1590/1519-6984.98660.

Plestenjak E, Neji M, Sinkovič L, Meglič V, Pipan B. 2025. Genomic insights into genetic diversity and seed coat color change in common bean composite populations. Front Plant Sci 15: 1523745. https://doi.org/10.3389/fpls.2024.1523745.

Purwanti E, Fauzi A. 2019. The morphological characteristics of *Phaseolus lunatus* L. in different areas of East Java, Indonesia. IOP Conf Ser Earth Environ Sci 276: 012017. https://doi.org/10.1088/1755-1315/276/1/012017.

Pusat Perlindungan Varietas Tanaman dan Perizinan Pertanian (PPVTPP). 2021. Panduan Penggunaan Aplikasi Pendaftaran Varietas Lokal dan Hasil Pemuliaan. PPVTPP, Indonesian Ministry of Agriculture, Jakarta. [Indonesian]

Puspita D, Palimbong S, Toy B, Notosoedarmo S. 2017. Identifikasi legum lokal di Pulau Timor yang berpotensi untuk pengembangan inovasi pangan lokal. Prosiding Seminar Nasional Pengembangan Sumber Daya Perdesaan dan Kearifan Lokal Berkelanjutan VII 324-335. Purwokerto, 17-18 November 2017. [Indonesian]

Quilichini TD, Gao P, Yu B, Bing D, Datla R, Fobert P, Xiang D. 2022. The seed coat's impact on crop performance in pea (*Pisum sativum* L.). Plants 11 (15): 2056. https://doi.org/10.3390/plants11152056.

Riptanti EW, Masyhuri M, Irham I, Suryantini A. 2019. The ability of dryland farmer households in achieving food security in the food-insecure areas of East Nusa Tenggara, Indonesia. AIMS Agric Food 5 (1): 30-45. https://doi.org/10.3934/agrfood.2020.1.30.

Sandrine MAD, Kevin KK, Clémence KL. 2020. Preliminary diversity assessment of lima beans (*Phaseolus lunatus*) cultivated in Côte d'Ivoire. Am J Plant Sci 11: 2059-2065. https://doi.org/10.4236/ajps.2020.1112145.

Sita K, Sehgal A, HanumanthaRao B, Nair RM, Vara Prasad PV, Kumar S, Gaur PM, Farooq M, Siddique KHM, Varshney RK, Nayyar H. 2017. Food legumes and rising temperatures: Effects, adaptive functional mechanisms specific to reproductive growth stage, and strategies to improve heat tolerance. Front Plant Sci 8: 1658. https://doi.org/10.3389/fpls.2017.01658.

Smýkal P, von Wettberg EJB, McPhee K. 2020. Legume genetics and biology: From Mendel's pea to legume genomics. Intl J Mol Sci 21 (9): 3336. https://doi.org/10.3390/ijms21093336.

Souza JDS, Rodrigues LN, Ferreira LT, Unêda-Trevisoli SH, Nascimento NFFD. 2023. Future-ready crops, genetic variability in lima bean seeds. Acta Sci Agron 46 (1): e66767. https://doi.org/10.4025/actasciagron.v46i1.66767.

Tudor EP, Lewandrowski W, Krauss S, Veneklaas EJ. 2024. Local adaptation to climate inferred from intraspecific variation in plant functional traits along a latitudinal gradient. Conserv Physiol 12 (1): coae018. https://doi.org/10.1093/conphys/coae018.

Wang Z, Zhao W, Huang Y, Zhao P, Yang K, Wan P, Chu L. 2023. Progress in adzuki bean seed coat color studies. Plants 12 (18): 3242. https://doi.org/10.3390/plants12183242.

Zhang H, Mascher M, Abbo S, Jayakodi M. 2022. Advancing grain legumes domestication and evolution studies with genomics. Plant Cell Physiol 63 (11): 1540-1553. https://doi.org/10.1093/pcp/pcac062.

Zhang JY. 2024. Legume evolution: From wild ancestors to modern crops. Legume Genom Genet 15 (3): 93-104. https://doi.org/10.5376/lgg.2024.15.0011.