

RESEARCH ARTICLE

Morpho-biochemical Characterization and Multivariate Assessment of Genetic Diversity in Rajmah (*Phaseolus vulgaris* L.) Genotypes from Northeast India

Bedashruti Saikia¹, Rupam Borgohain^{2*}, Gayatree Goswami², Jutika Das² and Anubrat Borah²

Abstract

The present study assessed morphological, yield, and biochemical diversity among forty rajmah (*Phaseolus vulgaris* L.) genotypes collected from five north eastern states of India, representing both vine and dwarf growth habits. Field experiments were conducted during the *rabi* seasons of 2022–23 and 2023–24 at Assam Agricultural University, Jorhat, using a randomized block design with three replications. Eighteen traits were studied, including ten morphological and eight biochemical parameters and were evaluated through univariate and multivariate analyses. Analysis of variance revealed highly significant genotypic differences for all traits, indicating substantial genetic variability. Vine types exhibited greater plant height, branching, pod number, and seed yield, whereas dwarf types recorded higher 100-seed weight and superior nutritional attributes, including fat, fibre, niacin, vitamin B₆, and folic acid. Correlation analysis showed strong positive associations between yield and key plant architecture traits. Principal component, cluster, and heatmap analyses grouped genotypes primarily by growth habit and regional origin. The outcomes of the study signify rich genetic diversity in the rajmah germplasm of Northeast India and identify key traits for simultaneous improvement of yield and nutritional quality.

Keywords: Biochemical traits, Genetic diversity, Multivariate analysis, Northeast India, *Phaseolus vulgaris*

¹Department of Plant Breeding and Genetics, Assam Agricultural University, Jorhat, Assam, 785013, India.

²Directorate of Research (Agri), Assam Agricultural University, Jorhat, Assam, 785013, India.

***Author for correspondence:**

rupam.borgohain@aau.ac.in

Received: 02/01/2026

Revised: 22/04/2026

Accepted: 24/04/2026

How to cite this article: Saikia B, R Borgohain, G Goswami, J Das and A Borah (2026) Morpho-biochemical Characterization and Multivariate Assessment of Genetic Diversity in Rajmah (*Phaseolus vulgaris* L.) Genotypes from Northeast India. *Indian J. Plant Genet. Resour.* 39(2): 43-52.

DOI: <https://doi.org/10.56093/ijpgr.v39i2.174765>

Introduction

The common bean (*Phaseolus vulgaris* L.) is among the most important grain legumes globally, serving as a major source of dietary protein, complex carbohydrates, fibre, and essential micronutrients such as iron and zinc (Broughton *et al.*, 2003; Blair *et al.*, 2010). Owing to its adaptability and nitrogen-fixing ability, it plays a vital role in food and nutritional security, particularly for smallholder farmers (Beebe *et al.*, 2010). In India, especially in the North-Eastern Hill Region, rajmah (dry beans) is a key component of local diets and cropping systems, cultivated across diverse altitudes and micro-climates (Singh *et al.*, 2014; Rai *et al.*, 2020). Despite its importance, indigenous rajmah germplasm of Northeast India remains poorly characterized for morphological, yield, and biochemical diversity.

Originating in the Americas (Kwak & Gepts, 2009), common bean was likely introduced into India through early trade and colonial routes during the 16th century (Indian Farming, 2017). Subsequent diversification under varied Himalayan micro-climates, local selection, and regional isolation fostered a secondary centre of diversity in Northeast India (Dutta *et al.*, 2015; Banoo *et al.*, 2020).

Previous studies have reported substantial variability in *P. vulgaris* for plant height, pod length, seed weight, and maturity period (Raturi *et al.*, 2017; Chahota *et al.*, 2020). Multivariate tools such as cluster and principal component analyses have been effective in identifying divergent groups and key discriminating traits (Andrade *et al.*, 2019; Girish *et al.*, 2021). Biochemical profiling for protein, fat, fibre, and vitamins has also provided valuable inputs for nutritional improvement in common bean (Celmeli *et al.*, 2018; Khan *et al.*, 2021). However, integrated multivariate evaluations combining morphological and biochemical traits remain limited for rajmah germplasm.

Multivariate approaches, including principal component analysis, hierarchical clustering, and correlation analysis, efficiently reveal trait interrelationships and major dimensions of variation (Jain *et al.*, 2019; Singh *et al.*, 2021) and have differentiated growth habits, regional adaptation, and nutritional quality groups in legumes (Oliveira *et al.*, 2019; Mwale *et al.*, 2020). Given the ecological and cultural diversity of Northeast India, rajmah genotypes are expected to exhibit wide variation shaped by environment and farmer-driven selection; however, systematic quantitative analyses remain scarce. Therefore, the present study aimed to assess variability among rajmah genotypes through integrated analyses of morphological, yield, and biochemical traits, to identify key traits for genetic improvement and nutritional enhancement.

Materials and Methods

Experimental site and plant materials

The study was conducted using 40 rajmah genotypes collected from five north-eastern states of India viz., Assam, Meghalaya, Mizoram, Arunachal Pradesh, and Nagaland represent both dwarf and vine growth habits (Supplementary Table 1). Genotypes were labelled using state-specific prefixes followed by numeric codes: AS (Assam), MR (Mizoram), NL (Nagaland), AR (Arunachal Pradesh), and ML (Meghalaya). Field experiments were carried out during the *rabi* seasons of 2022–23 and 2023–24 at the Instructional-cum-Research (ICR) Farm, Assam Agricultural University (AAU), Jorhat (26.75° N, 94.22° E; 87 m asl). The experiment followed a Randomized Block Design with three replications as per Fisher (1935). Each plot comprised two rows (3 m length, 45 cm × 20 cm spacing), and standard agronomic practices were adopted to ensure uniform crop growth.

Trait measurement and biochemical estimation

Eighteen traits were studied, including ten morphological (plant height, branches per plant, chlorophyll content index, days to 50% flowering, days to maturity, pods per plant, pod length, seeds per pod, 100-seed weight, and yield t/ha) and eight biochemical parameters (crude protein, fat, fibre, cooking time, riboflavin, niacin, vitamin B₆, and folic acid). Chlorophyll Content Index (CCI) was measured using a SPAD meter by taking non-destructive readings from the fully expanded upper canopy leaves of each genotype. For each replication, SPAD readings were recorded from 5 plants, and the values were averaged to obtain a representative chlorophyll estimate.

Crude protein, fat, and fibre contents were analysed following AOAC (2016) procedures. Vitamins were estimated spectrophotometrically using Lowry *et al.*, (1951) with modifications for legume seeds. Each analysis was performed in triplicate, and mean values were used for statistical computation.

Statistical analysis

Descriptive statistics (mean, range, SD, CV %) were computed to assess variability. Data normality was verified using Shapiro–Wilk tests and Q–Q plots. Analysis of variance was performed using the randomized block design model:

$$Y_{ij} = \mu + G_i + R_j + \epsilon_{ij}$$

Where Y_{ij} represents the observed value of the i^{th} genotype in the j^{th} replication, μ is the overall mean, G_i is the fixed effect of the i^{th} genotype, R_j is the random effect of the j^{th} replication, and ϵ_{ij} is the random error term associated with each observation.

Significant differences were separated using Tukey's HSD test ($\alpha = 0.05$), while non-parametric data were tested using Kruskal–Wallis and Dunn's post-hoc tests (Gomez & Gomez, 1984). A Welch's t-test was used to compare dwarf and vine growth types. All analyses were performed in R software (version 4.4.2; R Core Team, 2024) using *rstatix*, *ggplot2*, and *dplyr* packages (Wickham, 2016).

Multivariate and correlation analyses

Pearson's correlation coefficients were calculated to evaluate inter-trait associations and visualized through heatmaps (*corrplot*, *ggcorrplot*; Friendly, 2002). Principal Component Analysis (PCA) was carried out on standardized data (mean = 0, SD = 1) following Jolliffe (2002) to identify key traits contributing to variability. Eigenvalues > 1 were retained (Kaiser's criterion), and scree plots and biplots were generated to visualize trait

loadings and genotype grouping.

Hierarchical clustering was performed on principal component scores using Euclidean distance and Ward's linkage (Sneath & Sokal, 1973) to classify genotypes by plant type and region, thereby reducing multicollinearity among traits and ensuring that clustering was driven by the major, independent sources of phenotypic variation; regional variation in standardized mean trait values was visualized using heatmaps (Gu, 2016).

Data visualization

All graphical outputs, including PCA biplots, dendrograms, and heatmaps, were generated in R using the packages ggplot2, factoextra, ggdendro, and pheatmap for reproducibility and visual clarity.

Results and Discussion

Variation among genotypes

Analysis of variance revealed highly significant differences ($P < 0.001$) among the 40 rajmah genotypes for all morphological and biochemical traits, except seeds per pod, which also differed significantly ($P < 0.01$). This pronounced variability reflects the broad genetic base of germplasm collected from diverse agro-ecological regions of Northeast India and indicates substantial scope for selection and trait-based genetic improvement (Dutta *et al.*, 2015; Banoo *et al.*, 2020; de Paula *et al.*, 2024).

Descriptive statistics (Table 1) revealed appreciable variation for several agronomic traits, with plant height ranging from 30.00 to 192.33 cm ($CV = 15.75\%$), branches per plant from 10.67 to 30.33 ($CV = 7.59\%$), and pods per plant from 14.33 to 70.00 ($CV = 4.57\%$). Days to maturity varied from 110.33 to 132.00 days, indicating the presence of both early- and late-maturing genotypes. Seed yield ranged from 1.03 to 2.77 t/ha, demonstrating substantial scope for yield improvement. Among quality traits, 100-seed weight

varied widely from 20.67 to 105.33 g, reflecting marked diversity in seed size and market class preferences. Crude protein ranged from 14.03 to 23.47%, fat from 0.60 to 4.07%, crude fibre from 4.75 to 23.26%, and folic acid from 242.81 to 336.37 μg , indicating significant nutritional variability among genotypes.

Welch's t-test (data not shown) revealed clear phenotypic differentiation between growth habits. Vine types recorded significantly higher plant height, branching, pod number, seeds per pod, phenological duration, seed yield, and cooking time, reflecting vigorous growth and longer crop duration. In contrast, dwarf types exhibited significantly higher 100-seed weight and superior nutritional attributes, including crude fibre, fat, niacin, vitamin B₆, and folic acid contents. Traits such as chlorophyll content index, pod length, crude protein, and riboflavin did not differ significantly between growth habits, indicating selective rather than uniform effects of growth habit on trait expression. Similar growth habit-driven patterns have been reported in common bean germplasm (Gelaw, 2017; Shah *et al.*, 2020).

Following the detection of significant genotypic differences in ANOVA, Tukey's HSD post-hoc test was applied to identify specific pairwise differences among genotypes for each trait. Table 1 integrates descriptive statistics with trait desirability criteria and superior-performing genotypes identified through HSD comparisons. Vine genotypes were prominent for growth and yield-related traits, whereas several dwarf genotypes excelled in seed size, earliness, cooking quality, and nutritional attributes. Early-maturing genotypes were particularly relevant for short-duration and multiple cropping systems, while nutrient-dense entries may serve as valuable donors for biofortification. Overall, growth habit emerged as a major source of variability, and the evaluated genotypes displayed broad agronomic and nutritional diversity, consistent with earlier reports in common bean (Ahmad, 2018; de Paula *et al.*, 2024).

Table 1: Descriptive statistics, trait desirability, and superior genotype performance based on Tukey's HSD analysis for major traits of rajmah

Trait	Mean $\pm$ SD	Range (Min–Max)	CV (%)	HSD (0.05)	Nature of desirable performance	Desirable genotypes with mean values	Remarks
Plant height (cm)	83.23 $\pm$ 52.50	30–192.33	15.75	43.29	Vine: Taller plants (180–250cm) desirable Dwarf: Short plants (30–40cm) desirable	Vine: NL3 (192.33); AR7 (183.33); AR13 (182.00) Dwarf: AS2 (36.33); AS1 (35.33); NL5 (30.00)	Height determines plant architecture and suitability for vine or dwarf production systems.

Trait	Mean $\pm$ SD	Range (Min–Max)	CV (%)	HSD (0.05)	Nature of desirable performance	Desirable genotypes with mean values	Remarks
Branches/plant	19.44 $\pm$ 5.08	10.67–30.33	7.59	4.88	Higher desirable in both types	Vine: AR13 (30.33); NL3 (28.67); AR4 (27.67) Dwarf: AS7 (20.67); AS3 (20.33); AS4 (19.67)	Higher branching may increase reproductive sites and yield potential.
Days to maturity	120.32 $\pm$ 7.08	110.33–132	1.28	5.09	Lower desirable in both types	Vine: NL3 (128.00); AR11 (127.00); AR9 (127.00) Dwarf: MR1 (115.00); AS3 (114.33); AR12 (114.00)	Early maturity is useful for short-duration and multiple cropping systems.
Pods/plant	38.32 $\pm$ 14.73	14.33–70.0	4.57	5.78	Higher desirable in both types	Vine: AR7 (70.00); NL3 (68.33); AR6 (67.67) Dwarf: AS11 (37.33); AS3 (37.00); AS10 (36.00)	More pods per plant directly support higher productivity.
100-seed weight (g)	58.98 $\pm$ 16.93	20.67–105.33	2.21	4.30	Higher desirable (50–100 g)	Vine: AR2 (105.33); Dwarf: AR3 (88.33); AR12 (85.00)	Bold seeds are preferred for market value and consumer acceptance.
Seed yield (t ha ¹)	1.79 $\pm$ 0.47	1.03–2.77	12.47	0.74	Higher desirable (> 1.5 t/ha)	Vine: AR6 (2.77); AR7 (2.66); NL3 (2.37) Dwarf: AS7 (2.21); AS11 (1.82); AS3 (1.80)	High-yielding genotypes are useful for varietal improvement and cultivation.
Cooking time (min)	25.90 $\pm$ 3.32	20.33–36.67	4.89	4.19	Lower desirable (20–25 min)	Vine: AR2 (36.67) Dwarf: AS10(20.33); AS5 (20.67); ML3 (21.30)	Shorter cooking time indicates softer texture after cooking and reduces fuel consumption.
Crude protein (%)	17.70 $\pm$ 2.20	14.03–23.47	4.74	2.77	Higher desirable (> 20%)	Vine: AR9 (23.47) Dwarf: MR2 (21.43); MR4 (21.03); AS3 (20.50)	High protein genotypes are valuable for nutritional improvement.
Fat (%)	1.93 $\pm$ 0.93	0.6–4.07	9.04	0.58	Higher desirable (1.5–3%)	Vine: AR8 (3.83); AR11 (3.60) Dwarf: MR2 (4.07); AS6 (3.80)	Higher fat may improve energy value and nutritive quality.
Crude fiber (%)	8.19 $\pm$ 3.60	4.75–23.26	2.53	0.68	Moderate desirable (6–9%)	Vine: AR8(7.7); AR6(7.23); AR3(7.01) Dwarf: AS6(8.85); ML3(8.21);ML4(7.01)	Suitable fibre content contributes to dietary and health benefits.
Folic acid (μ g)	287.16 $\pm$ 22.43	242.81–336.37	1.30	12.36	Higher desirable (> 300 μ g)	Vine: ML3 (335.92) Dwarf: MR1 (336.37); AS2 (333.21)	High folic acid genotypes are useful for biofortification.

Correlation analysis

The Pearson correlation analysis revealed strong interrelationships among morpho-agronomic and quality traits (Fig. 1). Plant height showed strong positive correlations with branches per plant ($r = 0.75$), pods per plant ($r = 0.83$), days to flowering ($r = 0.89$), days to maturity ($r = 0.89$), and seed yield ($r = 0.73$), indicating that taller and more vigorous genotypes tended to be later maturing and more productive. Branches per plant were positively associated with chlorophyll content index ($r = 0.86$), pods per plant ($r = 0.64$), days to flowering ($r = 0.75$), days to maturity ($r = 0.71$), and seed yield ($r = 0.81$), confirming the importance of branching and canopy vigour in yield formation. Pods per plant also showed strong positive correlations with days to flowering ($r = 0.85$), days to maturity ($r = 0.84$), and seed yield ($r = 0.86$), highlighting pod number as a major yield component. Similar associations of pod number, branching, and plant vigour with seed yield have been reported in common bean by Raturi *et al.*, (2017) and Chahota *et al.*, (2020), supporting their use as indirect selection criteria. Seeds per pod had moderate positive associations with pod length ($r = 0.49$), days to flowering ($r = 0.48$), and yield ($r = 0.52$). In contrast, 100-seed weight was negatively associated with days to flowering ($r = -0.46$), indicating that larger-seeded genotypes tended to flower earlier, while its relationship with yield was weak ($r = 0.18$). Cooking time showed negative correlations with yield ($r = -0.18$), riboflavin ($r = -0.41$), and folic acid

($r = -0.31$), suggesting that faster-cooking genotypes may also possess favourable nutritional quality. Among biochemical traits, crude protein was positively correlated with fat ($r = 0.25$) and riboflavin ($r = 0.29$), whereas riboflavin, niacin, vitamin B₆, and folic acid showed moderate positive intercorrelations ($r = 0.22$ – 0.44), indicating coordinated accumulation of nutritional compounds. Overall, the results suggest that yield improvement in this germplasm can be effectively targeted through plant architecture traits, while nutritional traits remain partly independent, offering scope for simultaneous selection of high-yielding and nutrient-dense genotypes.

Principal component analysis (PCA)

Principal component analysis of 18 morphological and biochemical traits across 40 rajmah genotypes representing two growth habits and five Northeast Indian regions revealed that the first six principal components (eigenvalues > 1) explained 77.6% of the total variation, with PC1 (35.2%) and PC2 (11.2%) contributing most. The scree plot showed a sharp decline after PC2, indicating that the first two components captured the major variability (Fig. 2). PC1 was dominated by plant height, branches per plant, pods per plant, and seed yield, confirming their central role in productivity differentiation, whereas PC2 was primarily associated with 100-seed weight, crude protein, and cooking time, reflecting seed size and compositional traits. Nutritional traits such as fat, fibre, riboflavin, niacin, vitamin B₆, and folic acid loaded moderately on later components, indicating secondary contributions to genotype differentiation. These patterns are consistent with previous multivariate analyses in common bean germplasm (Cichy *et al.*, 2019).

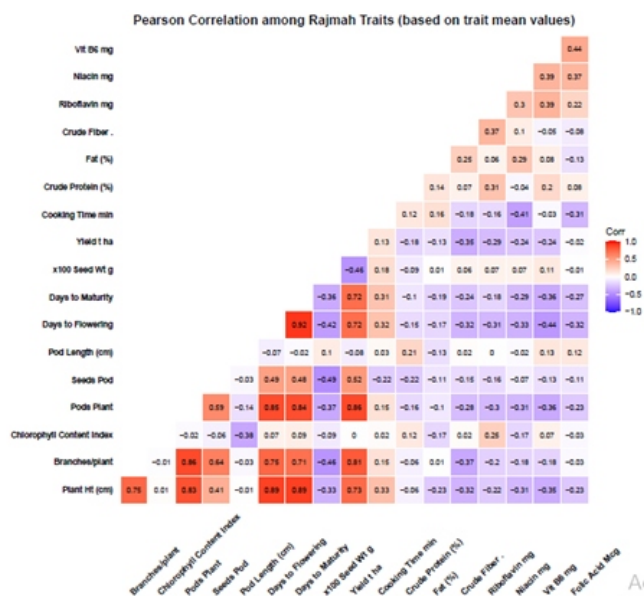


Fig. 1: Pearson correlation heatmap showing associations among 18 morphological and biochemical traits in 40 rajmah genotypes.

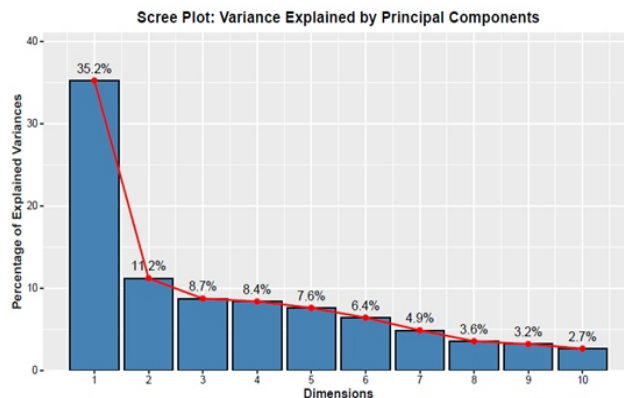


Fig. 2: Scree plot showing percentage variance explained by successive principal components.

In the PCA scatter plot based on plant type, vine genotypes clustered mainly along the positive axis of PC1 and were associated with taller plants, greater branching, and higher pod number, whereas dwarf genotypes grouped on the negative side of PC1, reflecting compact growth habit and comparatively lower yield potential. The limited overlap between groups indicated a few intermediate phenotypes (Fig. 3). Clear separation of growth habits in PCA space indicates that plant architecture contributes strongly to overall diversity, which agrees with observations of Andrade *et al.*, (2019) and Girish *et al.*, (2021), who also reported clustering of bean genotypes according to growth form and yield components. The PC1 vs PC3 projection further supported this separation, with vine genotypes showing a wider spread along PC3, suggesting greater within-group variation for

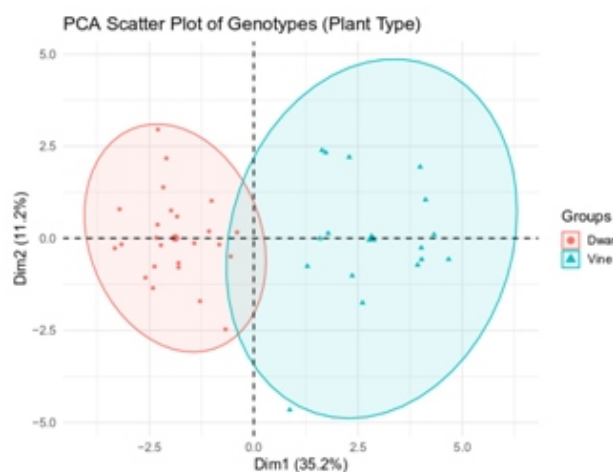


Fig. 3: PCA scatter plot of genotypes based on 18 traits (PC1 vs PC2), color-coded by plant type (Dwarf and Vine).

secondary traits such as maturity, seed quality, or nutritional attributes, while dwarf genotypes remained more compact (Fig. 4). Similarly, the PC2 vs PC3 plot showed partial differentiation between plant types, with vine genotypes tending toward higher positive PC2 scores and dwarf genotypes clustering nearer the origin or negative PC2 region, indicating that PC2 and PC3 captured additional but smaller sources of variation beyond the dominant growth-habit effect represented by PC1 (Fig. 5).

Regional PCA plots indicated that genotypes from Meghalaya and Arunachal Pradesh aligned more closely with yield-related traits on the positive side of PC1, whereas those from Assam and Nagaland



Fig. 4: PCA scatter plot of genotypes based on 18 traits (PC1 vs PC3), color-coded by plant type (Dwarf and Vine).

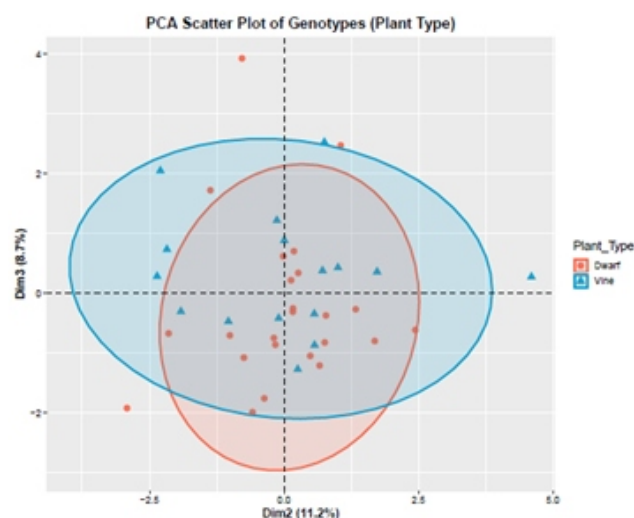


Fig. 5: PCA scatter plot of rajmah genotypes based on 18 traits (PC2 vs PC3), grouped by plant type (Dwarf and Vine).

clustered toward compact growth and earlier maturity, with Mizoram genotypes occupying intermediate positions (Supplementary Fig. 1). Collectively, these patterns highlight the influence of growth habit and regional adaptation on phenotypic diversity in rajmah and agree with earlier reports in common bean (Long *et al.*, 2020). PCA biplots illustrated that vectors for plant height, branches per plant, seeds per pod, pods per plant, and seed yield aligned with vine-type genotypes, while vectors for 100-seed weight and crude protein, fibre, fat, and vitamin traits aligned in the opposite direction, highlighting contrasting trait emphasis across growth habits (Fig. 6).

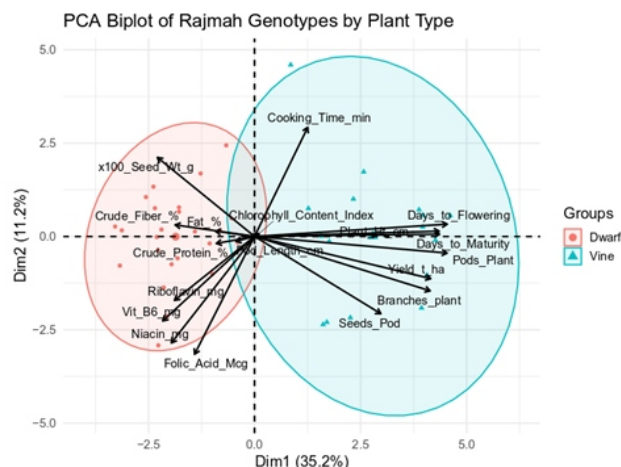


Fig. 6: PCA biplot of genotypes and trait vectors coloured by plant type.

Regional variation in standardised mean trait values

The hierarchical heatmap of standardised mean trait values (Fig. 7) displayed distinctive regional patterns: accessions from Arunachal Pradesh recorded the highest values for yield-related traits (height, branching, pods per plant and seed yield), Nagaland genotypes showed moderate yield but heavier seeds and longer cooking time, Mizoram genotypes were richer in nutritional traits (crude protein, fat, vitamins) and accessions from Assam and Meghalaya combined early maturity with lower yield traits. These results corroborate the PCA findings and confirm that regional origin strongly influences the trade-off between yield potential and nutritional composition in this germplasm collection.

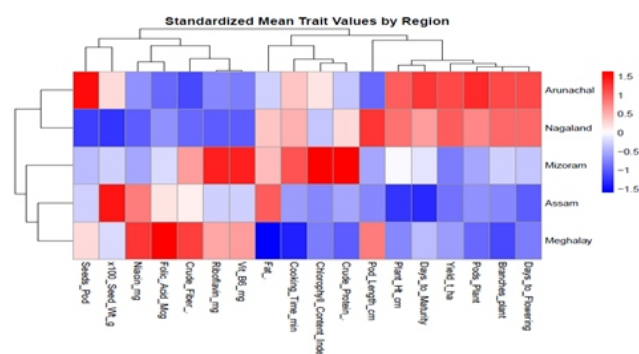


Fig. 7: Hierarchical heatmap showing standardized mean values of 18 morphological and biochemical traits across rajmah genotypes from five northeastern regions.

Cluster analysis and regional grouping of genotypes

Hierarchical cluster analysis based on PCA scores (Fig. 8) grouped the 40 rajmah genotypes into two major clusters broadly corresponding to dwarf and vine growth habits, with further sub-clusters reflecting regional origins. Accessions from Arunachal Pradesh

and Meghalaya formed one main cluster characterised by greater plant height and pod number, while accessions from Assam and Nagaland grouped separately, displaying compact growth and moderate yield. Mizoram accessions occupied intermediate positions between the two main clusters. This clustering confirms that growth habit is the principal determinant of phenotypic diversity in this collection, with regional adaptation contributing to intra-group variation. Similar clustering patterns based on morphological divergence have been reported in common bean germplasm from China, Ethiopia, and Brazil (Gelaw, 2017; Long *et al.*, 2020; de Paula *et al.*, 2024), supporting the robustness of the present findings.

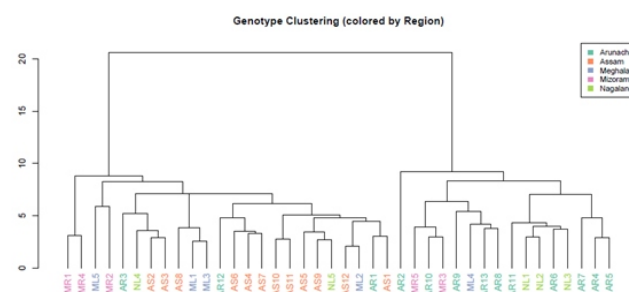


Fig. 8: Hierarchical dendrogram showing clustering of 40 rajmah genotypes by region based on multivariate trait data.

Breeding implications

The findings suggest that selection for improvement of rajmah germplasm in Northeast India can be effectively guided by a concise set of traits like plant height, branches per plant, pods per plant, 100-seed weight and crude protein. The clear separation of vine and dwarf types with distinct trait profiles indicates the potential for targeted crossing to combine high-yield and high-nutritional quality. In particular, vine-type accessions with strong yield component traits and dwarf-type accessions with superior seed quality can be selected as parents for developing genotypes combining both agronomic performance and nutritional value. The strong region-based differentiation also suggests the value of exploiting local adaptation in breeding strategies. Many of the underlying associations (e.g., yield component trait correlations) have been emphasized in recent meta-QTL and genomic studies of common bean (Arriagada *et al.*, 2023).

Conclusion

The present study demonstrated substantial morpho-biochemical diversity among rajmah genotypes

collected from Northeast India. Significant variation across agronomic, yield, and nutritional traits indicates strong potential for selection and genetic improvement. Multivariate analyses consistently revealed growth habit as the principal source of variation, with vine genotypes excelling in plant vigour and yield-related traits, while dwarf genotypes possessed superior seed size and several nutritional attributes. Regional grouping further suggested the influence of local adaptation on phenotypic expression. Traits such as plant height, branches per plant, pods per plant, 100-seed weight, and crude protein emerged as important selection criteria for breeding. The identified contrasting genotypes provide valuable parental resources for developing high-yielding, nutritionally enriched, and regionally adapted rajmah cultivars for the North-Eastern Hill Region and similar agro-ecologies.

Acknowledgements

The authors are grateful to the Assam Agricultural University, Jorhat, for providing the necessary facilities to conduct this study. Support received from the Department of Plant Breeding and Genetics and the Directorate of Research (Agri), AAU, during field and laboratory work is duly acknowledged.

References

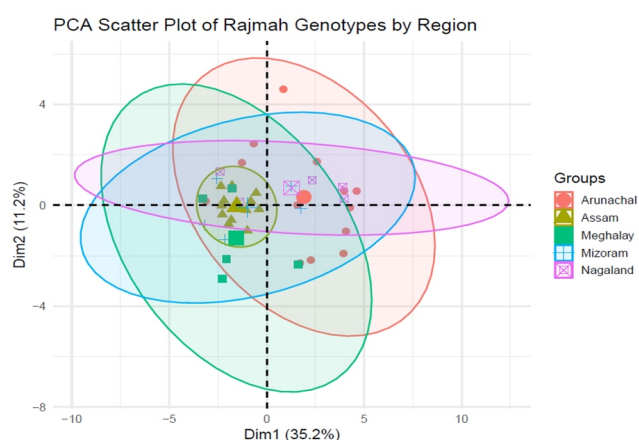
- Ahmad KMS (2018) Genetic diversity of common bean (*Phaseolus vulgaris* L.) cultivars from different origins revealed by microsatellite markers. *J. Adv. Biol. Biotechnol.* 17(4): 1-9.
- Andrade RA, LM Lima and EV Rodrigues (2019) Genetic diversity and structure in common bean based on morpho-agronomic traits. *Genet. Resour. Crop Evol.* 66(5): 1107-1121.
- AOAC (2016) *Official Methods of Analysis*, 20th edn. Association of Official Analytical Chemists International, Gaithersburg, Maryland, USA.
- Arriagada O, B Arévalo, RA Cabeza, B Carrasco and AR Schwember (2023) Meta-QTL analysis for yield components in common bean (*Phaseolus vulgaris* L.). *Plants* 12(1): 117.
- Banoo A, N Ahmed, SH Khan and SM Zargar (2020) Genetic diversity assessment in common bean (*Phaseolus vulgaris* L.) germplasm using agro-morphological traits. *Legume Res.* 43(5): 674-680.
- Beebe S, AV González and J Rengifo (2010) Research on trace minerals in the common bean. *Food Nutr Bull* 31(1 Suppl 1): 67-70.
- Blair MW, LF González, PM Kimani and L Butare (2010) Genetic diversity, inter-gene pool introgression and nutritional quality of common bean (*Phaseolus vulgaris* L.) from Central Africa. *Theor. Appl. Genet.* 121(2): 237-248.
- Broughton WJ, G Hernández, M Blair, S Beebe, P Gepts and J Vanderleyden (2003) Beans (*Phaseolus* spp.)-model food legumes. *Plant Soil* 252(1): 55-128.
- Celmeli T, H Sari, H Canci, D Sari, A Adak, T Eker and C Toker (2018) Nutritional content of common bean (*Phaseolus vulgaris* L.) landraces in comparison to modern varieties. *Agronomy* 8(9): 166.
- Chahota RK, TR Sharma and SK Sharma (2020) Genetic variability and trait association studies in rajmah (*Phaseolus vulgaris* L.) under mid-hill conditions of Himachal Pradesh. *Legume Res.* 43(6): 773-778.
- Cichy KA, JA Wiesinger, M Berry, S Nchimbi-Msolla, D Fourie, TG Porch, D Ambechew and PN Miklas (2019) Role of genotype and environment in determining cooking time of dry beans (*Phaseolus vulgaris* L.). *Legume Sci.* 1: e13.
- de Paula E, RN Almeida, TO Santos, JDN Souza Neto, EMS Riva-Souza, SCP Posse, MN Souza, AFF Madella de Oliveira, AC Santos Júnior and JO Santos (2024) Genetic diversity of common bean (*Phaseolus vulgaris* L.) landraces based on morphological traits and molecular markers. *Plants* 13(18): 2584.
- Dutta M, A Saikia, N Gogoi and JD Nath (2015) Genetic variability studies in pole-type rajmah (*Phaseolus vulgaris* L.) genotypes of north-eastern hill region of India. *Legume Res.* 38(5): 631-635.
- Fisher RA (1935) *The Design of Experiments*, 1st edn. Oliver and Boyd, Edinburgh, Scotland, 252 p.
- Friendly M (2002) Corrgrams; exploratory displays for correlation matrices. *The American Statistician* 56: 316-324.
- Gelaw YM (2017) Genetic diversity in speckled bean (*Phaseolus vulgaris* L.) genotypes in Ethiopia. *J. Plant Breed. Crop. Sci.* 9(11): 200-207.
- Girish M, G Muniswamy and S Ramesh (2021) Multivariate analysis for genetic divergence in rajmah (*Phaseolus vulgaris* L.) genotypes. *J. Hort. Sci.* 16(1): 47-54.
- Gomez KA and AA Gomez (1984) *Statistical Procedures for Agricultural Research*, 2nd edn. John Wiley and Sons, New York, USA, 680 p.
- Gu Z (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32(18): 2847-2849.
- Indian Farming (2017) Rajmah; an important pulse crop for hills. *Indian Farming* 67(9): 17-19.
- Jain A, P Sharma and V Kumar (2019) Morphological and biochemical characterization of common bean genotypes. *Indian J. Genet. Plant Breed.* 79(4): 737-744.
- Jolliffe IT (2002) *Principal Component Analysis*, 2nd edn. Springer-Verlag, New York, USA, 487 p.
- Khan S, A Singh and M Sharma (2021) Nutritional and biochemical evaluation of common bean genotypes for protein and micronutrient content. *Legume Res.* 44(6): 673-679.
- Kwak M and P Gepts (2009) Structure of genetic diversity in the two major gene pools of common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 118: 979-992.
- Long J, J Zhang, X Zhang, J Wu, H Chen, P Wang, Q Wang and C Du (2020) Genetic diversity of common bean germplasm resources based on morphological traits. *Front. Genet.* 11: 697.
- Lowry OH, NJ Rosebrough, AL Farr and RJ Randall (1951) Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* 193(1): 265-275.
- Mwale VM, JM Bokosi and MB Kwapata (2020) Morphological diversity of common bean landraces from Malawi. *Int. J. Plant Breed. Genet.* 14(1): 1-9.

- Oliveira JS, MF Pereira and PA de Carvalho (2019) Genetic diversity of common bean based on morphological and biochemical traits. *PesqAgropecu Bras* 54: e00013.
- R Core Team (2024) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.
- Rai N, M Rai and R Singh (2020) Characterization of common bean germplasm under North-Eastern hill conditions. *Indian J. Agric.Sci.* 90(2): 378-383.
- Raturi A, R Sharma and A Sharma (2017) Variability and correlation studies in common bean (*Phaseolus vulgaris* L.) genotypes. *Legume Res.* 40(3): 464-469.
- Shah AH, I Khaliq, G Bashir and S Jannat (2020) Genetic diversity in common bean germplasm of Jammu & Kashmir. *Indian J. Agric. Res.* 54(5): 629-634.
- Singh S, M Sood and A Sharma (2014) Genetic divergence and variability studies in rajmah (*Phaseolus vulgaris* L.). *Legume Res.* 37(4): 439-444.
- Singh Y, SK Yadav and S Sharma (2021) Multivariate genetic analysis of common bean genotypes. *Plant Genet. Resour.* 19(5): 432-440.
- Sneath PHA and RR Sokal (1973) *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. W.H. Freeman and Company, San Francisco, USA, 573 p.
- Wickham H (2016) *ggplot2: Elegant Graphics for Data Analysis*, 2nd edn. Springer International Publishing, Cham, Switzerland, 260 p.

Supplementary Table 1: Genotype designation, plant type, and state of origin of the rajmah genotypes evaluated

S. No.	Genotype designation	Plant type	State of origin
1.	ML1	Dwarf	Meghalaya
2.	ML2	Dwarf	Meghalaya
3.	ML3	Dwarf	Meghalaya
4.	ML4	Vine	Meghalaya
5.	ML5	Dwarf	Meghalaya
6.	NL1	Vine	Nagaland
7.	NL2	Vine	Nagaland
8.	NL3	Vine	Nagaland
9.	NL4	Dwarf	Nagaland
10.	NL5	Dwarf	Nagaland
11.	AR1	Dwarf	Arunachal Pradesh
12.	AR2	Vine	Arunachal Pradesh
13.	AR3	Dwarf	Arunachal Pradesh
14.	AR4	Vine	Arunachal Pradesh
15.	AR5	Vine	Arunachal Pradesh
16.	AR6	Vine	Arunachal Pradesh
17.	AR7	Vine	Arunachal Pradesh
18.	AR8	Vine	Arunachal Pradesh
19.	AR9	Vine	Arunachal Pradesh
20.	AR10	Vine	Arunachal Pradesh
21.	AR11	Vine	Arunachal Pradesh
22.	AR12	Dwarf	Arunachal Pradesh
23.	AR13	Vine	Arunachal Pradesh
24.	MR1	Dwarf	Mizoram
25.	MR2	Dwarf	Mizoram
26.	MR3	Vine	Mizoram

S. No.	Genotype designation	Plant type	State of origin
27.	MR4	Dwarf	Mizoram
28.	MR5	Vine	Mizoram
29.	AS1	Dwarf	Assam
30.	AS2	Dwarf	Assam
31.	AS3	Dwarf	Assam
32.	AS4	Dwarf	Assam
33.	AS5	Dwarf	Assam
34.	AS6	Dwarf	Assam
35.	AS7	Dwarf	Assam
36.	AS8	Dwarf	Assam
37.	AS9	Dwarf	Assam
38.	AS10	Dwarf	Assam
39.	AS11	Dwarf	Assam
40.	AS12	Dwarf	Assam

**Supplementary Fig. 1:** PCA scatter plot of genotypes grouped by region of origin.