Journal of Innovative Agriculture

Research Article

Multivariate trait-based framework for functional ideotyping in rice under irrigated conditions

Raveendran Muthurajan*, Williams Mohanavel, Ameena Premnath, Bharathi Ayyenar, Veera Ranjani Rajagopalan, Sudha Manickam

Department of Plant Biotechnology, Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore-641003, India.

*Correspondence
Raveendran Muthurajan
raveendrantnau@gmail.com

Volume: 12, Issue: 1, Pages: 9-17
DOI: https://doi.org/10.37446/jinagri/rsa/12.1.2025.9-17
Received: 2 September 2024 / Accepted: 14 February 2025 / Published: 31 March 2025

Background: A diverse rice genotypes were evaluated for fifteen morphological/growth, reproductive/yield-related, physiological, and phenological traits to uncover their genetic variability and trait associations that are critical for developing stress-resilient high-yielding genotypes.

Methods: Descriptive statistics, four-way Venn diagram, PCA biplot, correlation network, and Composite Performance Index (CPI) from standardized PC1 (grain yield-driven) with k-means clustering were employed to dissect variability and rank genotypes.

Results: Descriptive statistics revealed extensive phenotypic variation, particularly in grain yield per plant (4.95–45.5 g), plant height (PH) (68.1–168.0 cm), and total tillers /productive tillers, while intrinsic water use efficiency (iWUE) showed broad adaptation potential. A four-way Venn diagram of top-performing genotypes across source, sink, growth, and phenology categories highlighted limited functional overlap, emphasizing trait-specific excellence. Principal component analysis (PCA) biplot explained 39.9% of variation (PC1: 23.5%, PC2: 16.4%) and delineated three genotype clusters *viz.,* yield-oriented genotypes processing high total tillers, productive tillers, and spikelet fertility. Physiologically efficient genotypes having enhanced photosynthetic rate (PNET) and *i*WUE. Genotypes with large flag leaf area (FLA) and PH contributing for tall and biomass accumulation with lower yield efficiency. Correlation network analysis identified three interconnected clusters reproductive/yield-related, morphological/growth, and physiological. Plant height trait interconnects all these clusters in the correlation network. The Composite Performance Index (CPI), derived from standardized PC1 with grain yield as the directional driver, ranked genotypes continuously with k-mean clustering, discriminate the genotypes in to three clusters (Low, medium and high performing). Genotypes *viz.,* IRIS_313-10260, IRIS_313-9160, and IRIS_313-10609 identified as best performing genotypes.

Conclusion: Established a multivariate pipeline that integrates the morphological/growth, reproductive/yield-related, physiological, and phenological traits into an ideotype framework. No genotype excels across all modules, but top CPI performers (IRIS_313-10260, IRIS_313-9160, IRIS_313-10609) integrate complementary strengths, making them ideal core donors.

Keywords: *rice, genetic diversity, PCA biplot, correlation network, composite performance index, traits*

Introduction

The architecture of rice productivity is a multi-dimensional web of traits where source, sink, growth, and phenological traits interact through trade-offs and synergies, hence, single-trait breeding for productivity is less effective (Anwar et al., 2022; Li et al., 2019; Jyoti et al., 2024). Breeding for increased and sustainable rice productivity requires tools that summarise and prioritise complex, multivariate phenotypes. Multivariate methods such as principal component analysis (PCA) are widely used to reduce dimensionality and visualise the relationships among traits and genotypes, enabling breeders to identify trait axes relevant to yield and adaptation. PCA can be used for grouping genotypes and detecting major trait combinations in rice and other crops (Yano et al., 2019; Shi et al., 2021). Correlation network allows the examination of available trait architecture components in the dataset and identify the trait hubs that coordinate multiple functions. Correlation network has been used to reveal adaptive strategies under water-limited environments (Wang et al., 2023). Multi-trait indices have been recently used in selecting promising genotype (Ouattara et al., 2024). Accounting for the PCA, correlation network and multi-trait Composite Performance Index (CPI), the breeder can devise strategic crossing schemes by picking suitable genotypes to fulfil the breeding objectives that balances the yield trade-offs with other traits. Integrative multivariate analysis enables the identification of genotypes by combining favourable structural (source/sink) and physiological traits providing a practical path toward ideotype breeding in rice. In this study, we leverage descriptive statistics, PCA biplot, correlation network, and CPI ranking in a panel of 100 rice genotypes to (i) quantify phenotypic variation, (ii) reveal trait clusters and hubs, (iii) rank genotypes by multi-trait performance, and (iv) identify ideotype intersections for practical breeding use.

Materials and Methods

Plant material and crop husbandry

A total of 180 diverse rice accessions were raised at the Wetlands, Tamil Nadu Agricultural University, Coimbatore (latitude 11.01236°N, longitude 76.93559°E). Seeds were sown in raised nursery beds, and 21-day-old seedlings were transplanted to the main field at a spacing of 20 × 20 cm in 2 m² plots by following unreplicated single-plot field evaluation (observational design). Fertilizer was applied at the recommended rate of 150:50:50 kg NPK/ha, with basal application and two top-dressings following the TNAU Crop Production Guide 2020. All the cultural practices were practised as per TNAU Crop Production Guide 2020 (https://agritech.tnau.ac.in/pdf/AGRICULTURE.pdf).

Trait measurements

The fifteen traits were measured on five randomly tagged plants per genotype (unless otherwise specified) at appropriate phenological stages and grouped as follows:

Morphological/Growth Traits (plant structure and vegetative development)

PH (Plant height): Measured from the soil surface to the tip of the primary panicle at maturity and expressed in cm. TT (Total tillers): Total number of tillers per plant counted at maturity. PT (Productive tillers): Number of panicles bearing tillers per plant at maturity. FLA (Flag leaf area): Calculated from length x breadth x 0.75 correction factor and expressed in cm². VN (Total number of veins): Three flag leaf samples per genotype were collected at flowering; veins were counted manually and averaged. LBR (Grain length–breadth ratio): Length and breadth of 20 filled grains per plant were measured using a vernier caliper; LBR was computed as length ÷ breadth (no units).

Reproductive/Yield-Related Traits (direct influence on yield potential)

PL (Panicle length): Measured from the collar to the tip of the panicle (cm) on panicles collected for spikelet fertility assessment. TSN (Total spikelet number): Counted per panicle on five primary panicles per plant. SF (Spikelet fertility) was determined as (number of filled grains ÷ total spikelets) x 100 (%). GY (Grain yield per plant): Grains from primary and secondary panicles of five individually harvested plants were threshed, weighed separately, and averaged (g/plant).

Physiological Traits (resource use and functioning)

TRR (Transpiration rate, mmol m^{-2} s^{-1}), PNET (Photosynthetic rate, µmol m^{-2} s^{-1}), SC (Stomatal conductance, mmol m^{-2} s^{-1}): Recorded at flowering stage between 09:00 and 11:00 h on the middle portion of fully expanded flag leaves of three tagged plants per genotype using a portable photosynthesis system (ADC BioScientific Ltd., UK). *i*WUE (Intrinsic water use efficiency, µmol mol^{-1}) was calculated as PNET ÷ SC.

Phenological Trait (developmental timing)

DFF (Days to first flowering): Each plot was monitored twice daily (morning and evening); the day on which the first spikelet emerged from the boot leaf sheath was recorded. After recording morphological and physiological observations, tagged plants were harvested individually. Primary and secondary panicles were separated for yield and fertility assessments, and remaining plant parts were harvested in paper covers.

Statistical analyses

A sub set of 100 genotypes were selected based on the criteria that no missing data point for any genotype for any trait is available. All statistical analyses were performed using Python 3.9 with libraries including pandas, numpy, scikit-learn, matplotlib, seaborn, networkx, and venny4py. Phenotype values were standardized using StandardScaler prior to principal component analysis (PCA). PCA was conducted to reduce dimensionality, and the first two components were visualized using a biplot with trait loadings scaled for interpretability and 95% confidence ellipses were added per k-means cluster using eigenvalue decomposition. The Composite Performance Index (CPI) was computed as standardized PC1, with sign flipped if the GY loading was negative, ensuring higher values indicate superior performance. Correlation networks were constructed from Pearson coefficients with edges coloured by $|r|$ bins and styled by p-value thresholds ($p \leq 0.05$ dotted, $p \leq 0.01$ solid edge). Violin plots were generated using seaborn. Venn diagrams were created using venny4py with annotation of genotype within intersection regions.

Trait selection for Composite Performance Index (CPI)

The CPI methodology was followed for the present analysis. All measured traits (PH, TT, PT, FLA, VN, PL, TSN, SF, GY, LBR, TRR, PNET, SC, iWUE, DFF) were standardized (z-scores) and subjected to principal component analysis (PCA). Traits retained for CPI were identified by the method given below. Let a_j be the loading of trait j on the first principal component (PC1). Define the 75th percentile threshold of absolute PC1 loadings as $\tau = \text{quantile}(\{|a_j| : j=1,\ldots,p\}, 0.75)$. The preliminary set of PC1-important traits is $S=\{j: |a_j| \geq \tau\}$. From S retain only those traits that are positively correlated with grain yield $T = \{j \in S : \text{corr}(X_j, GY) > 0\}$. Where, corr is the Pearson correlation computed across genotypes. Finally, the CPI components set is $C = \{GY\} \cup \{i\text{WUE}\} \cup T$.

CPI formula (weighted with GY highest weight)

All traits in CCC were standardized to z-scores:

$$Z_{ij} = \frac{X_{ij} - \mu_j}{\sigma_{ij}}$$

where X_{ij} is the observed value of trait j for genotype i and μ_j, σ_j are the mean and standard deviation of trait j across genotypes.

We used a simple weighting scheme that gives grain yield (GY) the highest influence (weight $w_{GY}=2$) and assigns unit weight to all other CPI components. Let the CPI component weights be w_j (with $w_{GY}=2w$) and $w_j=1$ for all $j \in C \setminus \{GY\}$ Then the weighted CPI for genotype i is:

$$\text{CPI}_i = \frac{\sum_{j\in \text{C}} w_j Z_{ij}}{\sum_{j\in \text{C}} w_j}$$

which, for the typical case where $C=\{GY, i\text{WUE}, PL, PT, TT, DFF, LBR\}$, becomes

$$\text{CPI}_i = \frac{2Z_{i,\text{GY}} + Z_{i,i\text{WUE}} + Z_{i,\text{PL}} + Z_{i,\text{PT}} + Z_{i,\text{TT}} + Z_{i,\text{DFF}} + Z_{i,\text{LBR}}}{8}$$

Results

Phenotypic Variation of the measured traits

Descriptive analysis (Figure 1) depicts the distribution and variability of fifteen traits displaying a substantial genetic variation within the population for all the traits measured. Among morphological/growth traits, plant height (PH) showed a broad range of variation (68.1–168.0 cm; mean 116.7 cm), suggesting that the population is mixture of both dwarf and tall genotypes. Total tiller (TT; 6.0–33.0) and productive tiller (PT; 5.6–33.4) exhibited considerable variation displaying the availability of tillering potential in the population. Flag leaf area (FLA; 8.94–70.84 cm^2) displayed substantial spread, suggesting wide variation in source capacity. In contrast, vein number and length breadth ratio showed narrow distributions, indicating relative uniformity. Among reproductive/yield-related traits, grain yield per plant (GY; 4.95 to 45.5 g; mean 18.5 g), display a right-skewed distribution where a few genotypes contributed exceptionally high yields.

Total spikelet number (TSN; 22–212.6) also showed a wide spread, reflecting differences in reproductive capacity. Panicle length (PL) and spikelet fertility (SF) varied moderately (Figure 1). Among physiological traits, photosynthetic rate (PNET; 1.76–17.18 µmol m^{-2} s^{-1}) and Intrinsic water use efficiency (*i*WUE7; 43–64.85 µmol mol^{-1}) displayed substantial variation indicating that the availability of varied mechanisms for increasing photosynthetic efficiency and water conservation. Transpiration rate and stomatal conductance showed narrow variation in the population. Days to fifty percent flowering (DFF; phenological trait) had moderate variation, with weak links to other traits. Overall, the prominent variations in the reproductive/yield-related and physiological traits emphasis the rich genetic variability of the traits that can be explored for selecting superior, high-yielding, and resource-efficient genotypes in the breeding program.

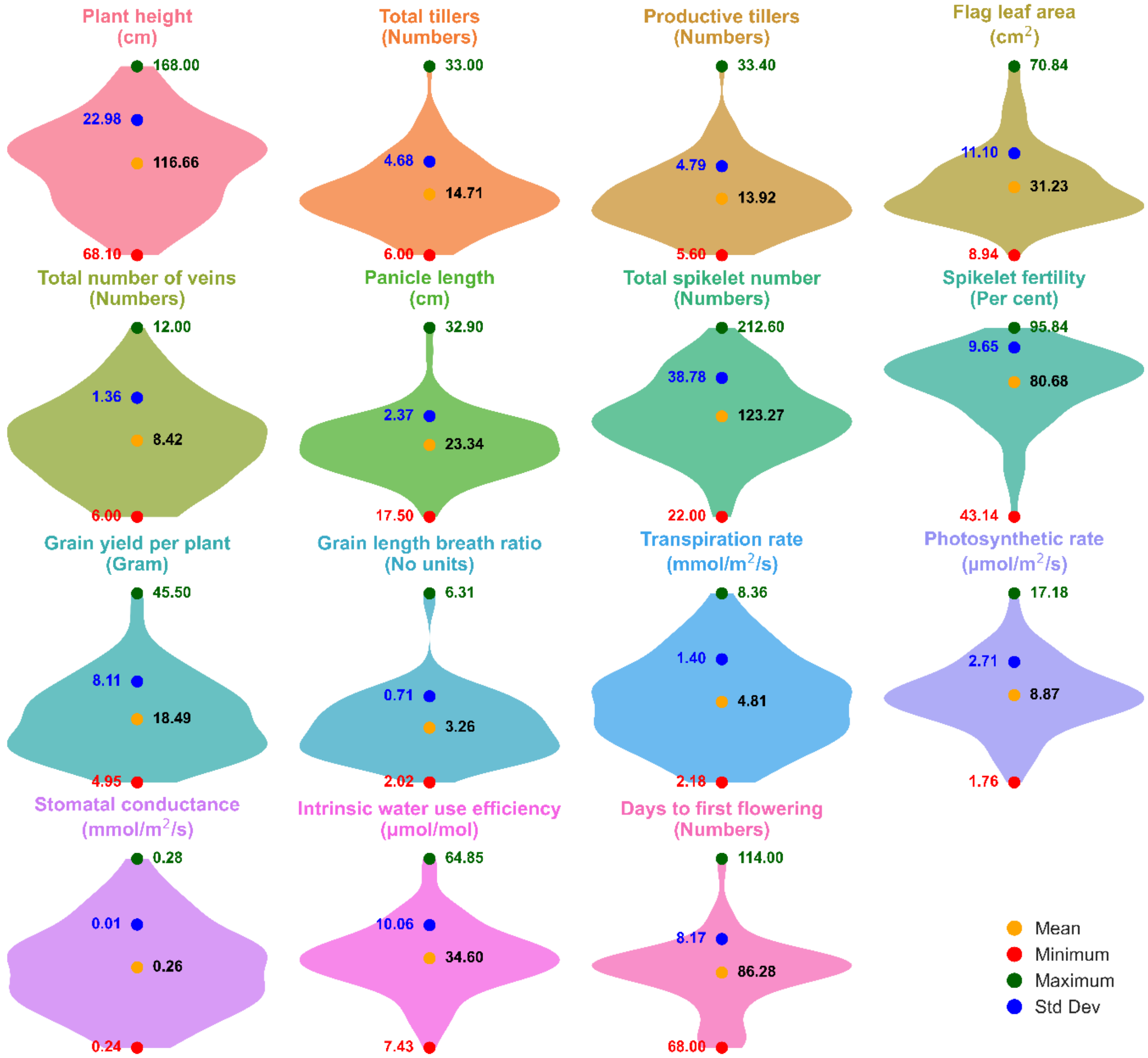

Figure 1. Violin plots showing the distribution of fifteen traits grouped into morphological/growth, reproductive/yield-related, physiological, and phenological categories. Large colored markers indicate minimum (red), maximum (dark green), mean (orange), and standard deviation (blue)

Overlap of Top-Performing Genotypes Across Functional Categories

The four-way Venn diagram (Figure 2) illustrates the overlap of the top 10 performing rice genotypes across four functional categories, Source (dominated by physiological traits: FLA, VN, PNET, TRR, SC, iWUE), Sink (reproductive/yield-related traits: PL, TSN, SF, LBR, GY), Growth (morphological/growth traits: PH, TT, PT), and Phenology (DFF), derived from their composite performance indices. A considerable number of genotypes were unique to individual categories, indicating strong trait-specific performance. The Growth category included seven unique genotypes (e.g., IRIS_313-10000, IRIS_313-10047), while Phenology had eight (e.g., IRIS_313-10016, IRIS_313-10260). Limited overlap of genotypes across three category and no overlap of genotypes across all the four categories was observed.

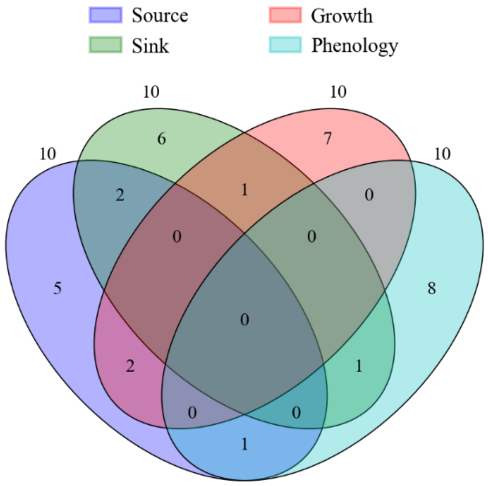

Figure 2. Four-way Venn diagram depicting overlaps of top 10 genotypes across source, sink, growth, and phenology categories identified by CPI ranking

Multivariate Relationships and Genotype Clustering

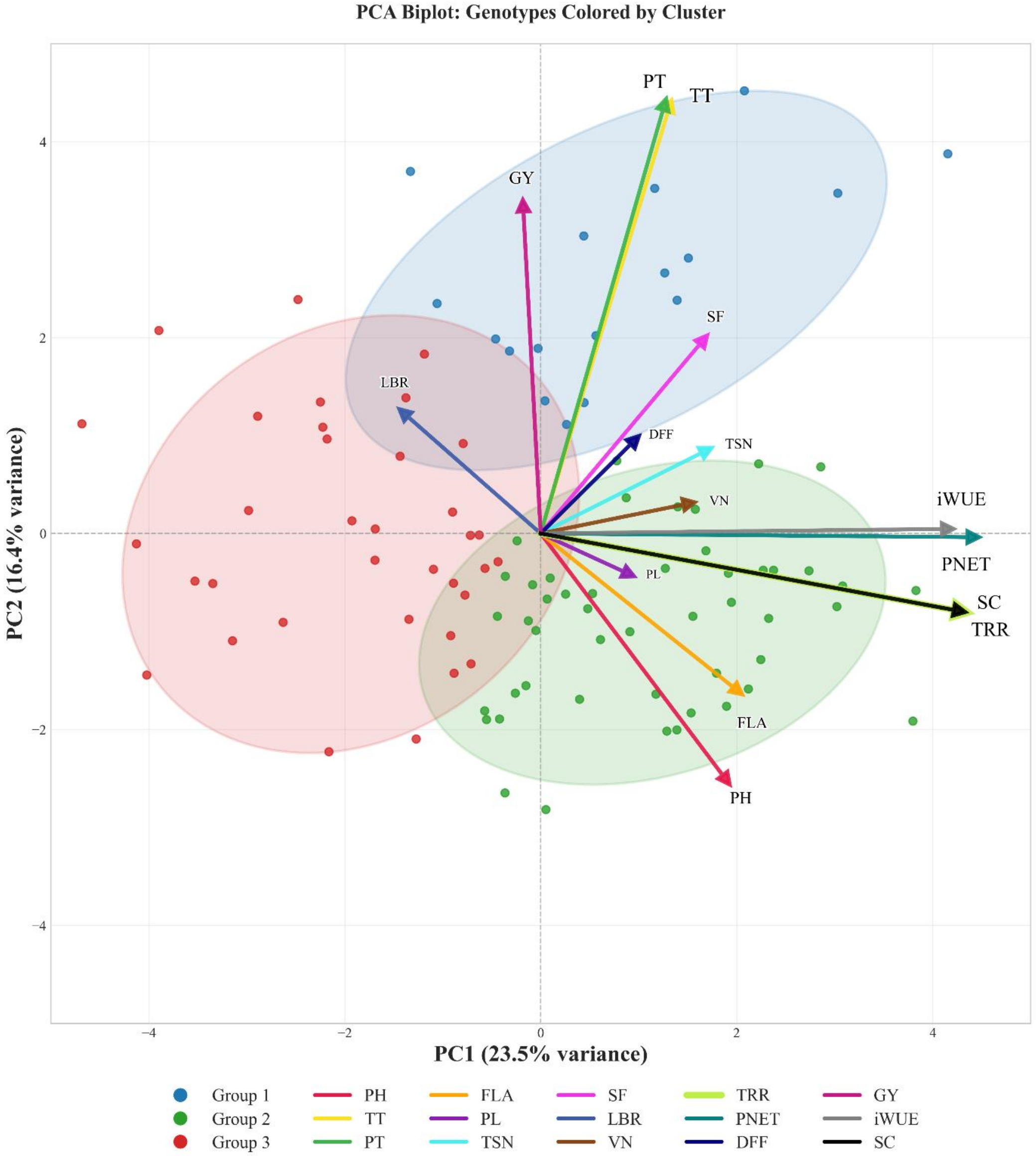

Figure 3. PCA biplot based on fifteen traits grouped with three clusters (yield-oriented: blue; physiologically efficient: green; vigorous, tall-statured: red) with 95% confidence ellipses

Computation of PCA (Figure 3) estimates that 39.9% of the total phenotypic variation was explained by PC1 (23.5%) and PC2 (16.4%) and cluster the genotypes into three distinct clusters. Group 1 (blue) comprising of yield-oriented genotypes, positioned upper-right in figure 3, showing strong positive loadings of reproductive/yield-related traits (TT, PT, SF, GY). The close alignment of GY, TT, and PT vectors indicates a strong positive correlation among these traits. The genotypes in this cluster will have efficient tiller-to-grain conversion leading to high-yield. The genotypes in Group 2 (green) are physiologically efficient which are located in the lower-right in the figure 3. These genotypes are associated with physiological traits (iWUE, PNET, SC, TRR), indicating a superior gas-exchange regulations and water-use efficiency. Group 3 (red) genotypes are vigorous, tall-statured situated left in figure 3 that align with morphological/growth traits (LBR, PH, FLA), but opposite to GY, TT, PT, suggesting that vegetative vigour was achieved at the cost of grain yield per plant. Longer vectors (e.g., TT, PT, GY, PNET, iWUE) indicate greater influence of these on the total phenotypic variation available in the population. Acute angles between TT, PT, and GY reflects the synergic effect of these traits while the opposite directions between PH/FLA and yield traits signify their trade-offs.

Trait Interconnections in Correlation Network

The correlation network (Figure 4) reveals three major interconnected clusters of traits. Reproductive/yield-related module comprises of GY, SF, TT, PT traits that are tightly linked *via* strong positive correlations reflecting their synergetic role in productivity. Morphological/growth module comprises of PH, FLA, PL, TSN traits that possibly connect linking canopy and panicle structures. Physiological module comprises of PNET, TRR, SC, iWUE, VN that are tightly connected, with SC, TRR, PNET as central nodes. PH serves as a linking hub between morphological/growth and physiological modules. DFF is weakly connected, confirming its independence.

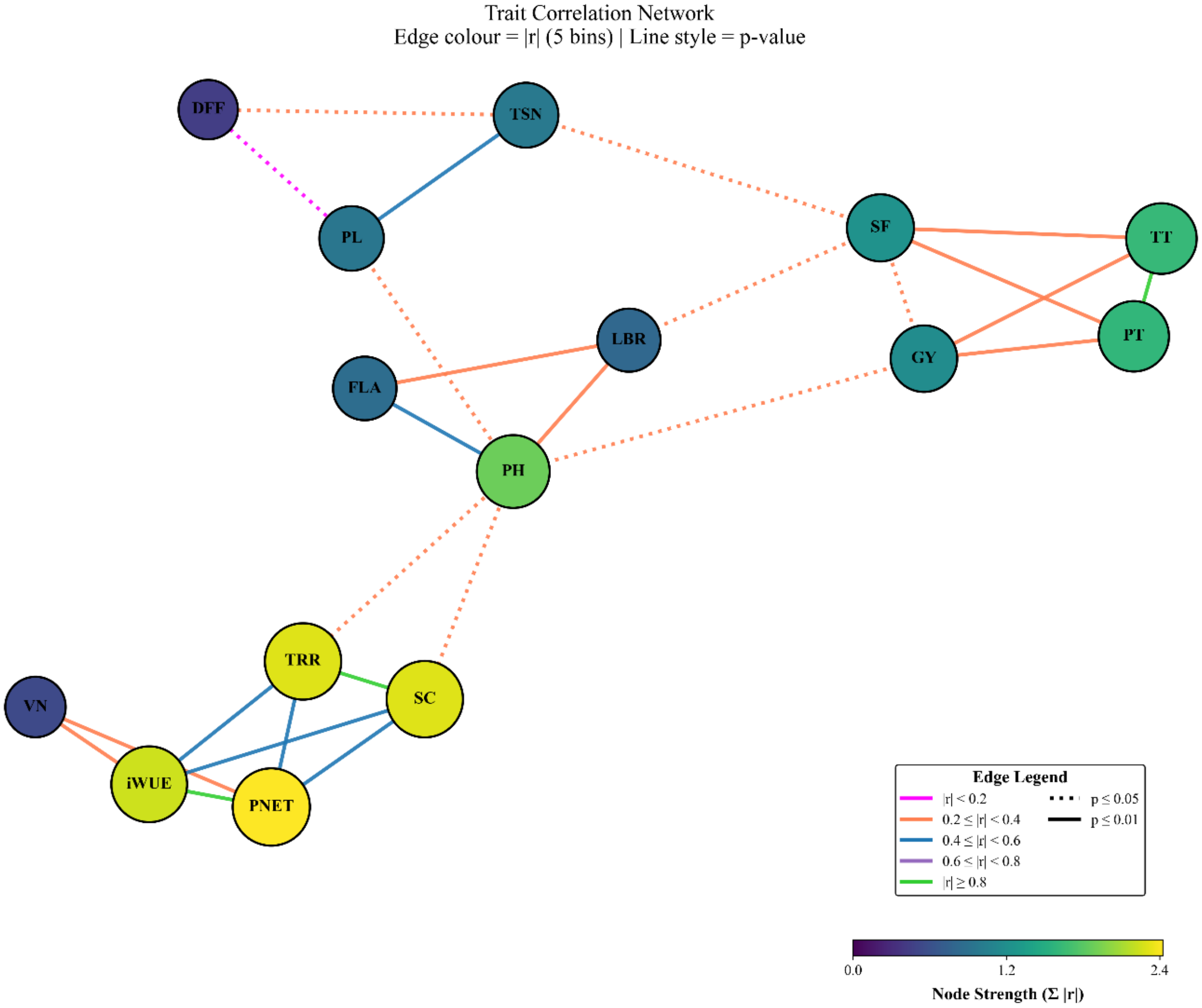

Figure 4. Correlation network illustrating inter-trait relationships. Three clusters: reproductive/yield-related (GY, TT, PT, SF), morphological/growth (PH, FLA, PL, TSN), physiological (PNET, SC, TRR, iWUE, VN). PH = central hub

Composite performance index and genotype ranking

The CPI dot plot (Figure 5) ranks all the genotypes by standardized PC1 with GY driving direction. The CPI ranking coupled with k-mean clustering identified three groups. Group 3 (dark green) had highest CPI score (0.24 to 4.68) top panel (figure 5; e.g., IRIS_313-10260, IRIS_313-9160). Group 2 (light coral) has lowest CPI score (-3.82 to 0.71). Group 1 (gold) was moderate (-4.15 to 1.33). The top 10 CPI genotypes are IRIS_313-10260, IRIS_313-9160, IRIS_313-10609, IRIS_313-9227, IRIS_313-11671, IRIS_313-11885, IRIS_313-11899, IRIS_313-10458, IRIS_313-10016 and IRIS_313-11638 having CPI score ranging from 4.68 to 2.62. This CPI integrates the synergetic traits of grain yield integrating the reproductive/yield-related and phenological traits without extreme morphological/growth dominance.

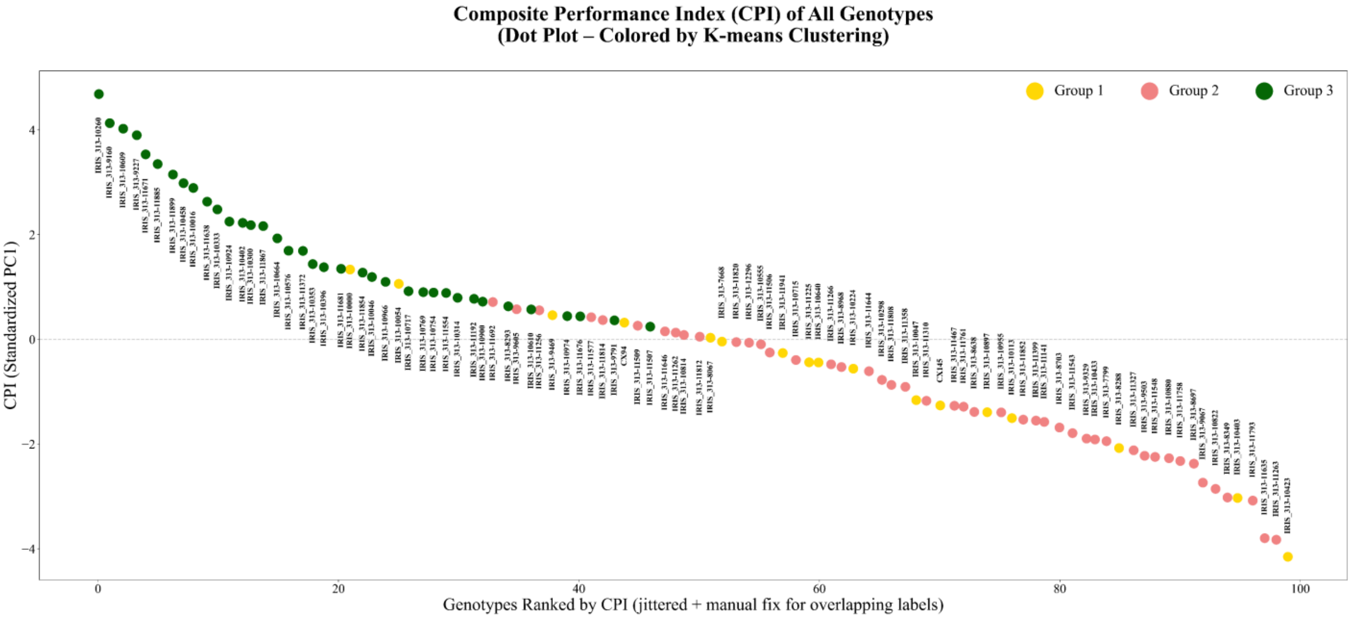

Figure 5. Composite Performance Index (CPI) dot plot with k-means clustering into three performance groups (gold, light coral, dark green). Genotypes ranked with jitter; labelled by ID

Discussion

The multivariate coherence across phenotypic diversity, shared association, and multi-trait interaction provides a robust framework for functional ideotyping in rice, highlighting a consistent structuring of genotype–trait relationships by trait groupings that are categorised in to morphological/growth, reproductive/yield-related, physiological, and phenological modules. Comprehensive selection accounting for these interactions enhances the efficiency of the breeding activity. Comprehensive multivariate approaches such as Multi-trait Genotype Ideotype Distance Index (MGIDI) and/or Composite Performance Index efficiently capture the multi-dimensional trait contributions to identify superior genotypes for yield or stress adaptive traits (Debnath, 2024; Habib et al., 2024). The right-skewed distribution of grain yield (Figure 1)) observed corresponds to a PCA cluster (Group 1; Figure 3) containing GY, TT, and PT) which defines yield-oriented genotypes. This finding was validated by strong reproductive/yield-related trait interconnections observed in the correlation networks that include SF, TT, PT, and GY.

These traits are crucial for achieving higher yield manifested through improved sink capacity and grain filling (Mahalakshmi et al., 2024). The Composite Performance Index (CPI) ranking identify the top-ranking genotypes with traits related to growth and phenology (e.g., TT, PT and DFF), prioritizing their importance in improving GY. This reinforces earlier reports emphasizing the coordination of phenological development and tiller dynamics for enhancing grain yield potential (Li et al., 2019). The physiological module involving PNET, *i*WUE, SC, and TRR are the characteristics of PCA Group 2 genotypes. These genotypes can be explored to develop drought tolerant genotypes using suitable breeding strategy. It is interesting to note that none of the genotypes in Group 2 are within the top 10 genotypes identified by CPI ranking and the genotypes in Group 2 has minimal overlap in shared associations (Figure 5) highlights that physiological efficiency traits primarily enhance performance under stress conditions rather than optimal conditions a scenario widely observed in stress tolerance studies (Habib et al., 2024).

The morphological/growth module, characterized by taller plant height (PH), high flag leaf area (FLA), and increased total spikelet number (TSN), was placed opposite to GY in PCA (Group 3). This antagonism lowers the CPI values of tall genotypes since high biomass accumulation without rationalizing the photosynthates to yield will reduce the harvest index (Cheng et al., 2022). PH act as a network hub that inversely related to GY suggests that taller canopies increase the transpiration load with the trade-off with grain yield under non-stress conditions.

This pattern aligns with earlier findings that tall rice plant architecture increases the water use but often reducing the harvest index (Jing et al., 2022). Venn diagram overlaps illustrating a limited shared associations among different trait module emphasizing that no single genotype excels in all the category (Table 1). High CPI is due to a balanced trait combination bridging growth, sink, and phenology modules. This balance strategy is critical for achieving wholistic performance, reinforcing the ideotype concept where complementary trait clusters driving the yield improvement (Habib et al., 2024).

Table 1. List of top 10 genotypes per functional category ranked by composite performance index (CPI)

Intersection	Count	Genotype
Source	10	IRIS_313-10403, IRIS_313-10423, IRIS_313-10822, IRIS_313-11263, IRIS_313-11635, IRIS_313-11758, IRIS_313-11793, IRIS_313-11852, IRIS_313-8349, IRIS_313-8697
Sink	10	IRIS_313-10433, IRIS_313-10822, IRIS_313-10897, IRIS_313-11327, IRIS_313-11554, IRIS_313-11681, IRIS_313-11793, IRIS_313-11941, IRIS_313-8293, IRIS_313-9329
Growth	10	IRIS_313-10000, IRIS_313-10047, IRIS_313-10113, IRIS_313-10224, IRIS_313-10403, IRIS_313-10423, IRIS_313-10897, IRIS_313-11225, IRIS_313-8067, IRIS_313-8288
Phenology	10	IRIS_313-10016, IRIS_313-10260, IRIS_313-10458, IRIS_313-10609, IRIS_313-10610, IRIS_313-11256, IRIS_313-11263, IRIS_313-11467, IRIS_313-11681, IRIS_313-12296
Source & Sink	2	IRIS_313-10822, IRIS_313-11793
Source & Growth	2	IRIS_313-10403, IRIS_313-10423
Source & Phenology	1	IRIS_313-11263
Sink & Growth	1	IRIS_313-10897
Sink & Phenology	1	IRIS_313-11681

Yield-oriented genotypes (Group 1) dominate optimal conditions *via* reproductive/yield-related synergy. Physiologically efficient genotypes (Group 2) represent the characteristics of a stress-tolerant, resource-efficient ideotypes characterized by higher intrinsic water-use efficiency and balanced gas-exchange regulation under irrigated conditions. Tall-statured genotypes (Group 3) need sink optimization. Top CPI genotypes are integrators. The complementary trait complexes identified through multivariate analyses can be pyramided *via* marker-assisted selection to develop ideotypes combining high yield and physiological efficiency.

Conclusion

This study established a multivariate pipeline that integrates the morphological/growth, reproductive/yield-related, physiological, and phenological traits into an ideotype framework. No genotype excels across all modules, but top CPI performers (IRIS_313-10260, IRIS_313-9160, IRIS_313-10609) integrate complementary strengths, making them ideal core donors. However, this study is based on a single-environment evaluation. GxE interactions could change trait correlations and CPI rankings. Future work should: (i) test genotypes across multiple environments and seasons; (ii) evaluate the genetic basis (QTL/GWAS) of identified hub traits; (iii) explore genomic-enabled multi-trait selection (genomic prediction) to accelerate selection cycles. Multi-trait genomic models have shown potential to increase prediction accuracy and selection efficiency.

Author contributions

RM: Conceptualized the study and manuscript revision. WM, AM and BA: Crop management and trait measurement. RR and SM analysed the data and drafted the manuscript.

Funding

No funding.

Ethics approval

Not applicable.

Competing Interests

The author declares no conflict of interest. The manuscript has not been submitted for publication in other journal.

References

Anwar, K., Joshi, R., Morales, A., Das, G., Yin, X., Anten, N. P., ... & Pareek, A. (2022). Genetic diversity reveals synergistic interaction between yield components could improve the sink size and yield in rice. *Food and Energy Security*, *11*(2), e334.

Cheng, X., Huang, Y., Tan, Y., Tan, L., Yin, J., & Zou, G. (2022). Potentially useful dwarfing or semi-dwarfing genes in rice breeding in addition to the *sd1* gene. *Rice (New York, N.Y.)*, *15*(1), 66.

Debnath, P., Chakma, K., Bhuiyan, M. S. U., Thapa, R., Pan, R., & Akhter, D. (2024). A novel multi trait genotype ideotype distance index (MGIDI) for genotype selection in plant breeding: application, prospects, and limitations. *Crop Design*, *3*(4), 100074.

Habib, M. A., Azam, M. G., Haque, M. A., Hassan, L., Khatun, M. S., Nayak, S., ... & Sarker, U. (2024). Climate-smart rice (*Oryza sativa* L.) genotypes identification using stability analysis, multi-trait selection index, and genotype-environment interaction at different irrigation regimes with adaptation to universal warming. *Scientific Reports*, *14*(1), 13836.

Jing, W., Wu, H., Gu, H., Xiao, Z., Wang, W., Zhang, W., ... & Zhang, H. (2022). Response of grain yield and water use efficiency to irrigation regimes during mid-season indica rice genotype improvement. *Agriculture*, *12*(10), 1647.

Jyoti, S. D., Singh, G., Pradhan, A. K., Tarpley, L., Septiningsih, E. M., & Talukder, S. K. (2024). Rice breeding for low input agriculture. *Frontiers in Plant Science*, *15*, 1408356.

Li, R., Li, M., Ashraf, U., Liu, S., & Zhang, J. (2019). Exploring the relationships between yield and yield-related traits for rice varieties released in China from 1978 to 2017. *Frontiers in Plant Science*, *10*, 543.

Mahalakshmi, P. J., Satish, Y., Raman, J. V., & Raju, M. R. B. (2024). Correlation and path coefficient analysis for yield and its component traits in advanced breeding lines of rice (*Oryza sativa* L.). *Journal of Advances in Biology & Biotechnology*, *27*(9), 769-777.

Ouattara, F., Agre, P. A., Adejumobi, I. I., Akoroda, M. O., Sorho, F., Ayolié, K., & Bhattacharjee, R. (2024). Multi-trait selection index for simultaneous selection of water yam (*Dioscorea alata* L.) genotypes. *Agronomy*, *14*(1), 128.

Shi, S., Wang, E., Li, C., Zhou, H., Cai, M., Cao, C., & Jiang, Y. (2021). Comprehensive evaluation of 17 qualities of 84 types of rice based on principal component analysis. *Foods*, *10*(11), 2883.

Wang, X., Ji, M., Zhang, Y., Zhang, L., Akram, M. A., Dong, L., ... & Deng, J. (2023). Plant trait networks reveal adaptation strategies in the drylands of China. *BMC Plant Biology*, *23*(1), 266.

Yano, K., Morinaka, Y., Wang, F., Huang, P., Takehara, S., Hirai, T., ... & Matsuoka, M. (2019). GWAS with principal component analysis identifies a gene comprehensively controlling rice architecture. *Proceedings of the National Academy of Sciences*, *116*(42), 21262-21267.