Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Phenotypic Discrimination of Routinely Identified Haploid and Diploid Sweet Corn Seedlings Using Multivariate Classification

Reynaldi Laurenze

reynaldilaurenze@unhas.ac.id

National Pingtung University of Science and Technology

Pearl Pei Chun Chang

National Pingtung University of Science and Technology

Muhammad Azrai

Hasanuddin University

Research Article

Keywords: Doubled haploid, diploid, haploid, haploid screening, multi trait phenotype, sweet corn

Posted Date: October 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10835072/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Early discrimination of haploid seedlings is an important step in doubled haploid (DH) breeding because haploid seedlings are typically identified among a much larger population of predominantly diploid progeny. Phenotypic traits expressed during early seedling development have been proposed as useful indicators of haploidy, but their combined discriminatory value across sweet corn breeding materials remains insufficiently characterized. This study evaluated the potential of integrated seedling phenotyping to discriminate seedlings routinely classified as haploid or diploid in sweet corn. Seedlings derived from controlled donor × inducer crosses were characterized for shoot growth, leaf development, root architecture, chlorophyll content, and stomatal traits. Principal component analysis (PCA), pearson correlation analysis, random forest, and linear discriminant analysis (LDA) were used to characterize multivariate phenotypic patterns, identify informative traits, and evaluate classification relative to the routine haploid/diploid classification. Haploid-classified seedlings generally exhibited lower plant height, shorter and narrower leaves, reduced root development, and smaller guard cells, whereas stomatal number showed the opposite pattern. Multivariate analysis indicated that shoot growth, root development, and stomatal characteristics contributed jointly to separation between the two routinely classified groups. Random forest identified leaf length, plant height, stomatal number, and leaf width among the more informative traits, while LDA provided multivariate discrimination between the two groups. These findings indicate that early seedling phenotypes contain information that can support preliminary discrimination of routinely identified haploid and diploid seedlings.

INTRODUCTION

Sweet corn (*Zea mays* L. var. saccharata) is a high-value horticultural commodity whose demand has steadily increased over the last two decades, due to the expansion of the fresh food sector, food processing industries, and rising consumer preferences for nutritious foods (Sidahmed et al. 2025; Rozi et al. 2023; Mushtaq et al. 2025). Unlike field corn, the economic value of sweet corn is mostly controlled by ear quality, sweetness, kernel texture, size uniformity, and visual appearance, therefore developing superior varieties is critical to preserving the commodity's competitiveness (Subaedah et al. 2021). In Indonesia, dry shelled corn production reached around 15.14 million tons in 2024, collected from approximately 2.55 million hectares, a 2.47% increase over the previous year (BPS 2025). Global grain production has surpassed 3.1 billion tons, with maize being a major contribution to the food, feed, and industrial sectors (Market Growth Report 2026). This growing demand has directly increased the strain on plant breeding efforts to create hybrid varieties with higher production, superior quality, and broad adaptability within shorter breeding cycles.

Despite rapid improvements in current breeding technologies, producing homozygous parental lines remains a major challenge. Under conventional breeding systems, pure lines are developed after six to eight generations of self pollination, which requires significant effort and financial expenditure (Taskin et al. 2023). Moreover, this challenge is exacerbated in sweet corn due to the fact that key quality traits are governed by numerous major genes, which need to be preserved throughout the line purification process

(Mandal et al. 2026; Ruanjaichon et al. 2021). Therefore, double haploid (DH) technology is one of the most significant innovations in modern plant breeding, allowing the production of fully homozygous lines in a single generation, thus shortening breeding cycles and improving selection efficiency (Zhao et al. 2025; Indu et al. 2023).

The DH technology is now a critical tool in modern maize breeding programmes, but its efficiency is limited by the low efficiency of haploid identification after induction (Chang et al. 2024; Khammona et al. 2024). Dermail et al (2024) and Guo et al (2026) reported that haploid identification is one of the most labor intensive stages of the entire DH production process because the proportion of haploid seeds generated by an inducer is relatively low compared to that of diploid seeds. Misclassification at this stage may result in diploid individuals being subjected to chromosome doubling, thereby increasing production costs, prolonging the selection process, and reducing the overall efficiency of breeding programs. Thus, haploid identification still remains a major bottleneck in the optimization of DH technology.

Several approaches have been developed to distinguish haploid from diploid seedlings in maize, including visual phenotypic markers, stomatal and guard cell characteristics, chloroplast related traits, and cytological or flow cytometric methods (Lopez et al. 2023; Chaikam et al. 2016; Chaikam et al. 2019; Souza et al. 2024). These approaches differ in their degree of specificity, measurement requirements, and suitability for early stage screening. Visual and phenotypic approaches can be implemented at relatively early developmental stages, whereas cytological and flow cytometric methods provide more direct evidence of ploidy but require specialized procedures and instrumentation. Consequently, phenotypic characteristics may provide useful information for preliminary discrimination of putative haploid seedlings, particularly when multiple traits are considered jointly rather than relying on a single morphological indicator (Uliana et al. 2022; Anirban et al. 2023).

Several types of fundamental biological processes controlling plant growth, such as cell division and expansion, stomatal development, ph otosynthetic activity, and root system formation, are known to be impacted by variations in ploidy level (Corneillie et al. 2019; Sabooni et al. 2022). As a result, these alterations might appear in physiological and morphological characteristics as early as the seedling stage. When compared to using the R1-nj marker alone, Chaikam et al (2017) found that seedling traits, such as radicle length, coleoptile length, and the number of lateral seminal roots, decreased misclassification rates. This suggests that seedling traits can act as a preliminary confirmation step before chromosome doubling. Additionally, Souza et al (2024) showed that guard cell length might disclose interactions between donor and inducer genotypes for this trait and differentiate between haploid and diploid seedlings during the early vegetative stage. These findings indicate that phenotypic traits have a great deal of potential as a basis for haploid identification. Nevertheless, their expression may be altered by genetic background, requiring a more thorough analytical approach.

However, previous studies have largely focused on identifying individual phenotypic differences within single donor populations or limited genetic backgrounds. Consequently, it remains unclear whether

these reported traits represent universally applicable indicators of ploidy or merely reflect genotype-specific responses. The practical value of phenotypic screening in breeding programs therefore depends not only on identifying discriminative traits but also on demonstrating that these traits remain robust across genetically diverse donor × inducer combinations. Although individual seedling traits have been investigated for haploid discrimination, less is known about how combinations of morphological, root, and stomatal traits jointly discriminate seedlings that have been routinely classified as haploid or diploid. In particular, it remains unclear whether a reduced subset of traits can retain discriminatory information, which is relevant to the practical burden of phenotyping

We hypothesized that haploid classified seedlings would exhibit coordinated changes in early seedling morphology, root development, and stomatal characteristics, and that multivariate combinations of these traits would provide stronger discrimination than individual traits alone.

Therefore, this study aimed to characterize early phenotypic differences between sweet corn seedlings routinely classified as haploid and diploid, evaluate their multivariate structure using PCA and correlation analysis, and assess the ability of LDA and Random Forest to discriminate between the two classification groups. A further objective was to identify a reduced set of informative traits that may support preliminary phenotypic screening.

MATERIAL AND METHODS

Sweet corn seed populations were generated through controlled crosses between multiple commercial hybrid donor genotypes (female parents) and a maternal haploid inducer line (male parent) under field conditions at the Experimental Farm, Faculty of Agriculture, Hasanuddin University, Makassar, Indonesia. The inclusion of multiple donor × inducer combinations was intentionally designed to evaluate whether early seedling phenotypic signatures associated with ploidy remained consistent across genetically diverse breeding populations rather than representing genotype-specific responses.

Artificial pollination was performed following the standard doubled haploid (DH) breeding protocol. Mature ears were harvested at physiological maturity, after which seeds were classified into putative haploid and diploid groups according to the routine identification procedure adopted in the breeding program. This classification served as the initial reference for subsequent phenotypic evaluation.

Classified seeds were sown under controlled greenhouse conditions using sterilized growth media to minimize environmental heterogeneity during early seedling development. All seedlings were maintained under identical irrigation, fertilization, and crop management practices throughout the experimental period. Environmental uniformity was maintained to ensure that phenotypic variation predominantly reflected differences in chromosome dosage rather than external growth conditions.

Early seedling phenotyping was conducted during the vegetative stage using a combination of morphological and physiological traits representing shoot growth, leaf development, root architecture, and epidermal cellular characteristics. Quantitative measurements included plant height, leaf length, leaf

width, leaf number, chlorophyll content, guard cell length, guard cell width, stomatal density, root length, and lateral root number. These traits were selected because previous studies have demonstrated that chromosome dosage influences cell expansion, tissue differentiation, and vegetative growth, making them biologically relevant indicators for early ploidy discrimination. Simultaneous evaluation of multiple trait categories was intended to capture coordinated phenotypic responses rather than relying on individual morphological characters.

Stomatal observations were performed using epidermal imprints obtained from the fully expanded leaf, followed by microscopic measurements. Chlorophyll content was measured using a portable chlorophyll meter according to the manufacturer's protocol.

Rather than evaluating individual traits independently, the analytical workflow was designed to identify a robust multivariate phenotypic signature associated with chromosome dosage. Consequently, complementary statistical and machine learning approaches were integrated to characterize trait relationships, reduce data dimensionality, prioritize informative variables, construct classification models, and validate phenotypic discrimination performance. All statistical analyses were performed using R software. The analytical framework consisted of descriptive statistics, inferential statistics, multivariate analyses, and machine learning approaches to comprehensively evaluate phenotypic differences between haploid and diploid seedlings.

1. Descriptive Statistics

Descriptive statistics were calculated for each morphological and physiological trait, including mean, standard deviation (SD), coefficient of variation (CV), and standard error (SE). Trait distributions were visualized using boxplots.

2. Principal Component Analysis

Principal Component Analysis (PCA) was performed using standardized trait values (z-score normalization). PCA was used to summarize multivariate phenotypic variation, visualize clustering patterns of haploid and diploid seedlings, identify traits contributing most strongly to overall variation. The contribution of each variable to the principal components was calculated and visualized using loading plots:

$$PC\kappa = \sum_{j=1}^{p} aj\kappa \, Xj$$

where $PC\kappa$ is the κ-th principal component, $aj\kappa$ is the loading coefficient of variable j, and Xj is the standardized value of the j-th phenotypic trait.

3. Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) was conducted to evaluate the discriminatory ability of combined morphological and physiological traits for classifying seedlings according to ploidy level. The standardized canonical coefficients were used to identify variables contributing most strongly to group discrimination. Model performance was evaluated using leave-one-out cross-validation:

$$LD = w^T x$$

where w is the vector of discriminant coefficients and x is the vector of standardized phenotypic variables

4. Random Forest Analysis

Random Forest classification was implemented to identify phenotypic traits with the highest predictive importance for ploidy classification:

$$\hat{Y} = \text{mode}\left\{T_1\left(X\right), T_2\left(X\right), \ldots, T_B\left(X\right)\right\}$$

where $T_i\left(X\right)$ represents the prediction from the i-th decision tree and B is the total number of trees in the forest.

Variable importance was quantified using the Mean Decrease in Accuracy and Mean Decrease in Gini Index.

$$VI_j = \frac{1}{B}\sum_{b=1}^{B} \Delta\, Error_{jb}$$

where VI_j is the importance score of variable j, and $\Delta\ Error_{jb}$ is the increase in classification error after permuting variable j in tree b. Traits with the highest importance values were considered candidate early phenotypic markers for haploid identification.

5. Reduced Traits Analysis Model

To evaluate the practical implications of measuring multiple phenotypic traits, a reduced trait analysis was performed. The full model included all evaluated phenotypic variables, whereas reduced models were constructed using traits identified as the most informative variables by the Random Forest analysis and supported by the multivariate analyses. Classification performance of the reduced models was compared with that of the full model using the same cross validation procedure.

6. Correlation Analysis

Relationships among morphological and physiological traits were evaluated using Pearson's correlation coefficient:

$$r_{xy} = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}\sqrt{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}}$$

where r_{xy} denotes the correlation coefficient between variables x and y, and n represents the number of observations. The resulting correlation matrix was visualized as a heatmap to investigate coordinated relationships among phenotypic traits and to identify biologically related trait modules associated with early seedling development.

The integration of unsupervised learning (PCA), supervised machine learning (Random Forest), and discriminant analysis (LDA) was intended to generate convergent evidence supporting the robustness and biological interpretability of the proposed early phenotypic screening framework.

RESULTS AND DISCUSSION

The present study demonstrated that haploid and diploid sweet corn seedlings exhibited clear phenotypic differences in several morphological and physiological traits during the early vegetative stage.

Compared with seedlings routinely classified as diploid, those routinely classified as haploid generally exhibited lower plant height, shorter and narrower leaves, fewer lateral roots, shorter roots, smaller guard cells, and a higher mean stomatal number (Table 1). Chlorophyll content showed a different pattern, with a higher mean value in the routinely classified haploid group. These results indicate that the two reference groups differed across multiple components of early seedling development, particularly shoot growth, root architecture, and stomatal characteristics.

Table 1
Descriptive statistics of phenotypic traits in routinely classified haploid and diploid sweet corn seedlings

Reference Classification	Trait	Mean	SD	SE	CV
Routinely classified diploid	Guard Length	36,70	4,38	0,89	11,92
Routinely classified haploid	Guard Length	27,17	5,02	1,02	18,48
Routinely classified diploid	Guard Width	15,77	2,45	0,50	15,56
Routinely classified haploid	Guard Width	12,02	3,03	0,62	25,16
Routinely classified diploid	Stomata	14,25	2,38	0,49	16,72
Routinely classified haploid	Stomata	21,88	2,80	0,57	12,81
Routinely classified diploid	Height	6,78	1,16	0,24	17,05
Routinely classified haploid	Height	3,18	1,09	0,22	34,13
Routinely classified diploid	Leaf Length	20,13	1,75	0,36	8,71
Routinely classified haploid	Leaf Length	10,60	4,05	0,83	38,17
Routinely classified diploid	Leaf Width	1,79	0,13	0,03	7,06
Routinely classified haploid	Leaf Width	1,24	0,25	0,05	20,19
Routinely classified diploid	Leaf Number	3,50	0,98	0,20	27,94
Routinely classified haploid	Leaf Number	3,04	1,12	0,23	36,89
Routinely classified diploid	Chlorophyll	5,08	0,46	0,09	8,99
Routinely classified haploid	Chlorophyll	5,81	0,94	0,19	16,21
Routinely classified diploid	Lateral Root	6,63	0,97	0,20	14,64
Routinely classified haploid	Lateral Root	4,50	1,32	0,27	29,31
Routinely classified diploid	Root Length	16,00	6,41	1,31	40,06
Routinely classified haploid	Root Length	8,95	3,75	0,77	41,96

The largest differences among the evaluated traits were observed in guard cell dimensions. Mean guard cell length was 36.70 in the routinely classified diploid group and 27.17 in the routinely classified haploid group, corresponding to an approximately 25.9% lower mean value in the latter group. Guard cell width showed a similar pattern, with mean values of 15.77 and 12.02, respectively. The consistent reduction in guard cell dimensions in the routinely classified haploid group is consistent with previous reports that associate ploidy level with differences in cell size (Yang et al. 2022; Marcondez et al. 2026; Sekiya et al. 2020). However, because chromosome number was not independently confirmed in the present study,

these differences are interpreted as phenotypic characteristics associated with the routine classification rather than as definitive cytological evidence of haploidy.

In contrast to guard cell size, stomatal number showed an opposite pattern. The routinely classified haploid group had a mean stomatal number of 21.88, compared with 14.25 in the routinely classified diploid group, corresponding to an approximately 53.5% higher mean value. The combination of smaller guard cells and a greater number of stomata in haploid plants represents a biologically important anatomical pattern. In general, changes in ploidy can influence epidermal cell size and the spatial constraints available on the leaf surface. One possible explanation is plants with smaller epidermal cells may accommodate a greater number of stomata per unit area. The relationship between guard cell length and stomatal number observed in this study therefore exhibited a complementary pattern, in which haploid plants were characterized by smaller guard cells but a greater number of stomata, whereas diploid plants exhibited larger guard cells but fewer stomata (Ribeiro et al. 2020; Maqbool et al. 2020).

Pronounced differences were also observed in vegetative growth traits. Diploid plants reached a mean initial height of 6.78 cm, whereas haploid plants reached only 3.18 cm, representing an approximately 53.1% reduction in haploid plants. A similar pattern was observed for leaf length, with mean values of 20.13 cm in diploid plants and 10.60 cm in haploid plants, corresponding to an approximately 47.3% reduction in haploids. Leaf width was also greater in diploid plants, reaching 1.79 cm compared with 1.24 cm in haploid plants. Collectively, these results indicate that haploid plants exhibited lower vegetative vigor than diploid plants. This phenotype may be associated with reduced genomic dosage, which can constrain meristematic activity, cell division, cell expansion, and biomass accumulation (Boerman et al. 2020).

Root system traits exhibited a pattern consistent with the reduced vigor of haploid plants. Diploid plants had a mean lateral root number of 6.63, compared with 4.50 in haploid plants, while mean root lengths reached 16.00 and 8.95 cm, respectively. Thus, root length in haploid plants was approximately 44.1% lower than that in diploid plants. These differences indicate that the effects of ploidy level extend beyond shoot growth and also influence the development of belowground organs. Root development depends strongly on meristematic activity, cell division, cell elongation, and lateral root formation. Therefore, changes in genomic dosage may affect the overall architecture of the root system. Previous studies in maize have similarly demonstrated that seedling traits, including root length and lateral root number, can contribute to the discrimination of haploid and diploid plants during early development (Baleroni et al. 2021; Marcondez et al. 2026).

Figure 3 illustrates the distribution of the evaluated traits according to the routine haploid/diploid classification. The differences were apparent across several traits rather than being restricted to a single variable. The routinely classified diploid group generally showed higher values for guard-cell dimensions, plant height, leaf length, leaf width, lateral root number, and root length, whereas the routinely classified

haploid group showed a higher mean stomatal number. This coordinated pattern supports the use of a multivariate approach because several traits may provide complementary information for discrimination.

Correlation analysis showed that early seedling traits did not vary independently but instead formed coordinated developmental modules associated with vegetative growth and epidermal differentiation. The dendrogram presented in Fig. 4 further revealed the clustering of traits according to their correlation patterns. The first cluster comprised vegetative growth and root-related traits, particularly height, leaf length, leaf width, root length, and lateral root, which exhibited relatively strong positive correlations with one another. The second cluster included guard length and guard width, which were strongly correlated with each other and positively associated with vegetative growth traits. In contrast, stomatal number formed a distinct cluster opposite to most growth-related traits because of its strong negative correlations. Chlorophyll appeared relatively isolated from the main clusters, reflecting its generally weak correlations with the other traits. This clustering structure demonstrates that the evaluated traits were not entirely independent but instead formed several phenotypic groups representing distinct, yet interconnected, biological processes.

The strongest positive correlation was observed between height and leaf length ($r = 0.93$). This very strong positive relationship indicates that seedlings with greater plant height tended to develop longer leaves. Physiologically, this association reflects the functional integration between the root system as an organ responsible for water and nutrient acquisition and the shoot as the primary photosynthetic organ. A more extensively developed root system can enhance resource availability to support leaf expansion and stem growth, whereas a larger shoot increases the plant's demand for water and nutrients, thereby potentially stimulating further root development (Magsayo et al. 2025; Ge et al. 2025).

The positive correlation between leaf length and leaf width ($r = 0.76$) indicates that leaf expansion occurred in a relatively coordinated manner. Similarly, the positive correlations of lateral root with leaf width ($r = 0.72$), leaf length ($r = 0.70$), and height ($r = 0.67$) indicate that lateral root development was closely associated with overall plant vigor. Lateral roots increase the soil volume explored by the root system and enhance the capacity for water and nutrient acquisition, whereas larger leaves may increase photosynthate production, thereby providing greater carbon resources to support root system development. This relationship reflects coordinated biomass allocation between belowground and aboveground organs (Zhu et al. 2024). Meanwhile, leaf number exhibited positive correlations with most growth-related traits, particularly lateral root ($r = 0.62$), suggesting that greater leaf initiation capacity tended to occur in parallel with increased root system complexity.

Stomata exhibited the most contrasting correlation pattern among the evaluated morphological traits. This parameter was negatively correlated with most growth-related traits, particularly leaf length ($r = -0.74$), leaf width ($r = -0.73$), root length ($r = -0.72$), and height ($r = -0.72$). This pattern indicates that higher stomatal number tended to be associated with smaller vegetative organs. The negative relationships may reflect a trade-off between stomatal characteristics and organ expansion, particularly leaf expansion. As the leaf lamina expands, the available surface area for stomatal distribution

increases, potentially resulting in a lower stomatal number per unit area. Conversely, smaller leaves may maintain relatively higher stomatal number (Fan et al. 2025; Bilgici et al. 2026).

In contrast to stomatal number, chlorophyll exhibited relatively weak associations with most morphological traits. Negative correlations were observed with height (r = − 0.51), leaf length (r = − 0.52), and root length (r = − 0.35), whereas its correlation with stomatal number was weak (r = 0.21). This pattern suggests that chlorophyll content did not consistently covary with overall plant size or stomatal variation. Physiologically, this may be explained by the fact that chlorophyll content primarily reflects the capacity for light energy capture, whereas stomatal traits are more directly associated with the regulation of gas diffusion and plant water status (Martínez-Barradas et al. 2026). Although both are essential components of photosynthesis, they are regulated by distinct physiological mechanisms and therefore do not necessarily exhibit strong covariation.

The PCA summarized the multidimensional phenotypic variation and showed separation between seedlings according to their routine haploid and diploid classification (Fig. 5). Because donor specific effects were not independently modeled, this separation should be interpreted as an overall phenotypic association with the routine haploid or diploid classification rather than definitive evidence that ploidy accounted for more variation than genetic background. Biologically, this pattern may be associated with the consequences of altered genomic dosage on plant growth and development. The reduced genomic content of haploid individuals may affect the dosage of genes regulating cell division and expansion, thereby influencing organ size, vegetative vigor, and overall plant development (Shi et al. 2021). In maize, haploid diploid differences have been consistently characterized by reductions in plant height, leaf length, leaf area, biomass, and other growth-related traits in haploid plants (Molenaar et al. 2019; Fuente et al. 2019). Therefore, the separation of individuals along the primary axis in the present study likely reflects the cumulative effects of phenotypic changes associated with differences in genomic dosage rather than random phenotypic variation alone. The broader distribution of haploid individuals compared with diploids further indicates greater phenotypic heterogeneity within the haploid group. This variation may arise because the expression of haploid effects is influenced by genetic background, such that the magnitude of vigor reduction is not necessarily uniform among individuals or germplasm backgrounds.

The contributions of height, leaf length, leaf width, root length, and lateral root to the same phenotypic gradient can be explained by developmental coordination between the shoot and root systems. Changes in genomic dosage under haploid conditions may systematically affect cell proliferation and expansion, resulting in simultaneous consequences for both aboveground and belowground organs. Accordingly, plants with more restricted shoot development may also exhibit reduced root development because the growth of both systems depends on coordinated cell division, cell expansion, and assimilate allocation (Azrai et al. 2025). The literature on haploid maize phenotypes supports this interpretation, showing that haploid plants exhibit reduced vigor from the seedling stage through maturity, including differences in plant size and biomass (Vanous et al. 2019; Yavuz et al. 2025). Thus, the simultaneous contribution of shoot and root related traits in the present study indicates that ploidy differences are expressed as an integrated growth syndrome rather than as isolated changes in individual organs.

In contrast, the stronger contribution of chlorophyll to Dim2 indicates that physiological variation did not fully follow the primary growth related gradient. Biologically, this may occur because chlorophyll content is determined by the interaction of multiple processes, including chloroplast development, pigment biosynthesis, nutritional status, leaf age, and environmental conditions (Martins et al. 2023; Mayta et al. 2019). Consequently, differences in plant size associated with ploidy level do not necessarily result in proportional changes in chlorophyll content. Previous studies on haploid identification have also demonstrated that certain physiological traits may be influenced by growth conditions, including light intensity, resulting in patterns of physiological variation that differ from those of morphological traits (Ulyanov et al. 2022; Ma'arif et al. 2025). Therefore, the positioning of chlorophyll along a distinct dimension may indicate that haploid diploid differentiation involves several biological processes that are not fully coordinated along a single axis. Whereas some traits primarily reflect genomic dosage and organ growth, physiological traits may more strongly reflect the regulation of metabolism and photosynthetic function.

To determine which phenotypic variables consistently contributed to ploidy discrimination, Random Forest classification was employed to rank trait importance. Figure 6 shows that leaf length exhibited the highest predictive importance, followed by plant height, stomatal density, and leaf width. Collectively, these variables represented complementary aspects of shoot expansion and epidermal development. The importance ranking suggests that discrimination was associated primarily with a combination of shoot growth and stomatal characteristics rather than a single phenotypic variable. Because variable importance was derived from a relatively small dataset, these rankings should be considered exploratory and require confirmation in larger independent populations. The dominance of these two traits is likely not attributable solely to differences in genomic dosage but also to their nature as integrative traits that reflect the cumulative outcomes of plant growth processes. Leaf length is determined by the coordinated interaction of cell division rate, cell expansion, leaf developmental duration, and assimilate availability during the vegetative phase (Li et al. 2022). Similarly, plant height represents the cumulative outcome of sequential organ growth and elongation processes (Baute et al. 2015). Consequently, relatively small changes in multiple growth processes may accumulate and manifest as substantial phenotypic differences in leaf length and plant height. In maize, haploid plants are known to exhibit shorter leaves and a smaller plant habit from the seedling stage, while previous studies have demonstrated that differences in seedling traits, including organ length and root development, can contribute to the discrimination of haploid and diploid individuals (Dermail et al. 2024).

The high importance of stomatal number across both importance metrics also has a biological basis distinct from that of the growth related traits. Stomata serve as major regulators of gas exchange and water loss. Thus, stomatal traits represent the interaction between epidermal development and leaf physiological function (Xue et al. 2024; Pan et al. 2024). Differences in stomatal number may influence CO_2 exchange capacity and transpiration regulation, whereas changes in stomatal size and distribution may affect water use efficiency and plant carbon balance (Nguyen et al. 2023). The high predictive importance of stomatal number therefore suggests that haploid diploid discrimination is not determined

solely by differences in plant size but also involves anatomical features associated with the regulation of gas exchange and water relations.

Guard length also exhibited relatively high importance because it represents a more specific cellular trait than macroscopic morphological characteristics. Differences in guard cell length may be influenced by patterns of cell expansion, epidermal developmental regulation, and the activity of factors controlling stomatal formation. Consequently, guard length may capture developmental differences at the cellular level that are not necessarily reflected by plant height or leaf length. Silva et al. (2019) demonstrated that guard cell length could consistently distinguish haploid and diploid individuals, although its discriminatory performance was influenced by germplasm source and genetic combination. The present model therefore appears to discriminate haploid individuals not only on the basis of their smaller overall stature but also through microscopic differences in leaf epidermal structure.

Leaf width and root length exhibited intermediate importance, which may be explained by their greater sensitivity to multiple sources of variation compared with leaf length and height. Leaf width is determined not only by longitudinal growth but also by lateral expansion, leaf shape, leaf position, and environmental conditions. Root length, meanwhile, is highly responsive to water availability, nutrient supply, substrate structure, and other physical environmental conditions. Consequently, although both traits may differ between haploid and diploid individuals, environmental variation may increase within-group variability and reduce their ability to function as strong standalone predictors. Nevertheless, their contribution remains biologically important because root development is an integral component of seedling vigor. Previous studies have shown that root traits, including radicle length and lateral root number, can contribute to the discrimination of haploid and diploid individuals. However, their effectiveness may depend on growth conditions and the population under evaluation.

The relatively low importance of chlorophyll, and particularly leaf number, also provides biologically meaningful information. Chlorophyll content is strongly influenced by leaf age, light intensity, nitrogen status, nutritional conditions, and the surrounding microenvironment, such that its variation does not necessarily directly reflect ploidy status. In maize, chlorophyll content and chloroplast related traits may also respond substantially to light conditions, potentially resulting in greater environmental noise. Leaf number, in contrast, is more closely associated with developmental stage and the timing of observation. When individuals are evaluated at relatively similar developmental stages, leaf number may not yet exhibit sufficient variation to serve as a major discriminator between haploid and diploid plants.

Table 2
Classification performance of the full 10-trait LDA model and the reduced trait LDA model

Number of Traits	Traits included	Accuracy (%)	Sensitivity (%)	Specificity (%)
2	Leaf Length, Height	97,92	95,83	100,00
3	Leaf Length, Height, Stomata	97,92	100,00	95,83
4	Leaf Length, Height, Stomata, Leaf Width	100,00	100,00	100,00
5	Leaf Length, Height, Stomata, Leaf Width, Guard Length	100,00	100,00	100,00
6	Leaf Length, Height, Stomata, Leaf Width, Guard Length, Chlorophyll	100,00	100,00	100,00
7	Leaf Length, Height, Stomata, Leaf Width, Guard Length, Chlorophyll, Root Length	100,00	100,00	100,00
8	Leaf Length, Height, Stomata, Leaf Width, Guard Length, Chlorophyll, Root Length, Lateral Root	97,92	100,00	95,83
9	Leaf Length, Height, Stomata, Leaf Width, Guard Length, Chlorophyll, Root Length, Lateral Root, Guard Width	97,92	100,00	95,83
10	Leaf Length, Height, Stomata, Leaf Width, Guard Length, Chlorophyll, Root Length, Lateral Root, Guard Width, Leaf Number	100,00	100,00	100,00

The classification performance of the LDA models further demonstrated that the phenotypic traits identified by random forest as important variables were sufficient to distinguish the routinely classified haploid and diploid seedlings with high accuracy. When traits were sequentially added according to their random forest importance (Table 2), the two-trait model consisting of leaf length and height achieved an overall accuracy of 97.92%, with sensitivity and specificity of 95.83% and 100.00%, respectively. Addition of stomata as the third trait maintained an accuracy of 97.92% but increased sensitivity to 100.00%, while specificity decreased slightly to 95.83%. Importantly, inclusion of leaf width as the fourth trait resulted in 100.00% accuracy, sensitivity, and specificity. Thus, the combination of leaf length, height, stomata, and leaf width provided complete agreement with the routine classification in the evaluated dataset.

the improvement obtained from two to four traits indicates that classification performance was not determined simply by the number of variables, but by the information contributed by specific phenotypic traits. leaf length and height provided the initial discriminatory information, whereas the addition of stomata and leaf width appeared to resolve the remaining overlap between the two classification groups. this finding is consistent with the random forest variable importance analysis, which identified these traits among the most informative variables. the concordance between the variable-selection and LDA

results therefore provides independent support for the relevance of these four traits as a reduced phenotypic set for seedling classification.

Interestingly, increasing the number of traits beyond the four-trait model did not improve classification performance. The five-, six-, and seven-trait models maintained 100% accuracy, sensitivity, and specificity, indicating that the additional traits did not provide substantial discriminatory information beyond the first four variables. This pattern suggests that additional phenotypic variables do not necessarily increase classification accuracy and may introduce variation that affects the decision boundary of the LDA model. Therefore, the four-trait model represents a more parsimonious solution because it achieved the same maximum classification performance as the full model while requiring less phenotypic measurement.

The broader distribution of haploid individuals observed in Fig. 7 may be associated with greater variation in their responses to environmental conditions, particularly because haploid individuals may possess different genetic buffering capacities compared with diploids (Muenchrath et al. 2023). In diploid plants, the presence of two alleles at many loci may provide a degree of redundancy or compensatory capacity when one allele exhibits suboptimal expression or function (Dermail et al. 2024; Khammona et al. 2025). Under haploid conditions, however, allelic variation is directly exposed at the phenotypic level because only a single allelic copy is present at each locus (Walters et al. 2013). Consequently, genetic differences among individuals may be more readily translated into phenotypic differences. This may result in a broader range of responses among haploid individuals exposed to the same environmental conditions. Thus, the greater heterogeneity of haploids in LDA space should not necessarily be attributed solely to haploid frailty, but may also reflect reduced genetic buffering capacity against variation in genotype and environmental response.

Another factor that may explain differences in phenotypic structure is variation in developmental stability and canalization. Canalization refers to the ability of a genotype to maintain a relatively stable phenotype despite environmental variation or developmental perturbations (Rünneburger & Le Rouzic 2016). If canalization capacity differs between haploid and diploid individuals, the same environmental conditions may generate different levels of phenotypic variation between the two groups. Individuals experiencing relatively small differences in germination, resource availability, or microclimatic conditions may progressively diverge in their developmental trajectories during subsequent growth. The accumulation of these developmental deviations may ultimately result in a broader phenotypic distribution in one group. Therefore, the broader distribution of haploid individuals may be interpreted as an indication of differences in developmental stability among individuals, without necessarily assuming that all observed variation results from reduced vigor.

CONCLUSION

The present study showed that sweet corn seedlings routinely classified as haploid and diploid differed in a coordinated set of early phenotypic traits, particularly shoot growth, leaf development, root

architecture, and stomatal characteristics. Multivariate analyses indicated that these traits jointly contained discriminatory information, while random forest identified leaf length, plant height, stomatal number, and leaf width among the more informative variables. LDA further demonstrated that the combined phenotypic profile could discriminate the two routine classification groups. These findings support the potential use of selected early seedling traits as a preliminary and complementary screening approach in sweet corn DH breeding.

Declarations

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The plant materials used in this study were obtained from controlled crosses conducted by the authors as part of the sweet corn breeding and research program. The collection, cultivation, and use of the plant materials complied with applicable institutional, local, and national guidelines and regulations. No protected or endangered plant species were involved, and no specific permits or licences were required for the use of these cultivated plant materials.

CONSENT FOR PUBLICATION

Not applicable.

COMPETING INTERESTS

The authors declare that they have no competing interests.

FUNDING

The authors received no specific funding for this work.

Author Contribution

R.L. conceived the study, conducted the data analysis, interpreted the results, and wrote the main manuscript text. P.P.C.C. contributed to the experimental design, interpretation of the results, and revision of the manuscript. M.A. contributed to the study conceptualization, supervision, interpretation of the findings, and critical revision of the manuscript. R.L., P.P.C.C., and M.A. reviewed and approved the final version of the manuscript.

DATA AVAILABILITY

The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. Anirban A, Hong HT, O'Hare TJ. Profiling and Quantification of Anthocyanins in Purple-Pericarp Sweetcorn and Purple-Pericarp Maize. Molecules. 2023;28(6):2665. https://doi.org/10.3390/molecules28062665.
2. Azrai M, Aqil M, Efendi R, Zainuddin B, Syam'un E, Anshori MF, Riadi M, Priyanto SB, Andayani NN, Yasin M, Laurenze R, Syam A. Canopy architectural and physiological traits for optimizing maize yield under high planting density. Front Sustain Food Syst. 2025;9:1469305. https://doi.org/10.3389/fsufs.2025.1469305.
3. [BPS] Statistic Indonesia. Luas Panen dan Produksi Jagung di Indonesia 2024. Jakarta; 2025.
4. Baleroni A, Pelozo A, Kamphorst. Samuel & Pedroza Carneiro, José & Rossi, Robson & Scapim, Carlos. (2021). Identification of haploids and diploids in maize using seedling traits and flow cytometry. Crop Breeding and Applied Biotechnology. 21. 10.1590/1984-70332021v21n4a54
5. Baute J, Herman D, Coppens F. Correlation analysis of the transcriptome of growing leaves with mature leaf parameters in a maize RIL population. Genome Biol. 2015;16:168. https://doi.org/10.1186/s13059-015-0735-9.
6. Bilgici M, Ebrahimi E, de Miranda LP, Lira S, Borras L, Young T, Yavuz R, Moore KJ, Dixon P, Lübberstedt T. Trends in stomatal number and size in maize hybrids representing 100 years of long-term breeding for yield. Front Plant Sci. 2026;17:1829321. https://doi.org/10.3389/fpls.2026.1829321.
7. Boerman NA, Frei UK, Lübberstedt T. Impact of Spontaneous Haploid Genome Doubling in Maize Breeding. Plants (Basel Switzerland). 2020;9(3):369. https://doi.org/10.3390/plants9030369.
8. Brestic M, Zivcak M, Hauptvogel P, Misheva S, Kocheva K, Yang X, Li X, Allakhverdiev SI. Wheat Plant Selection for High Yields Entailed Improvement Anatomical and Biochemical Traits Including Tolerance to Non-optimal Temperature Conditions. Photosynth Res. 2018;136(2):136–255. 10.1007/s11120-018-0486-z.
9. Chaikam V, Martinez L, Melchinger AE, Schipprack W, Boddupalli PM. Development and validation of red root marker-based haploid inducers in maize. Crop Sci. 2016;56(4):1678–88. https://doi.org/10.2135/cropsci2015.10.0653.
10. Chaikam V, Lopez LA, Martinez L, Burgueño J, Boddupalli PM. Identification of in vivo induced maternal haploids in maize using seedling traits. Euphytica: Neth J plant Breed. 2017;213:177. https://doi.org/10.1007/s10681-017-1968-3.
11. Chaikam V, Molenaar W, Melchinger AE. Doubled haploid technology for line development in maize: technical advances and prospects. Theor Appl Genet. 2019;132:3227–43. https://doi.org/10.1007/s00122-019-03433-x.

12. Chang Y, Tang H, Wang S, Li X, Huang P, Zhang J, Wang K, Yan Y, Ye X. Efficient induction and rapid identification of haploid grains in tetraploid wheat by editing genes TtMTL and pyramiding anthocyanin markers. Front. Plant Sci. 2024;15:1346364. 10.3389/fpls.2024.1346364.
13. Corneillie S, De Storme N, Van Acker R, Fangel JU, De Bruyne M, De Rycke R, Geelen D, Willats WGT, Vanholme B, Boerjan W. Polyploidy Affects Plant Growth and Alters Cell Wall Composition. Plant Physiol. 2019;179(1):74–87. https://doi.org/10.1104/pp.18.00967.
14. Darmawan M. (2014) Induksi Pembungaan di Luar Musim pada Tanaman Jeruk Keprok (*Citrus reticulata*). Master's thesis, Retrieved from https://pdfs.semanticscholar.org/2602/55a00e355dd178d668f08d4f6ae75ddd5868.pdf
15. Dermail A, Mitchell M, Foster T, Fakude M, Chen Y-R, Suriharn K, Frei UK, Lübberstedt T. Haploid identification in maize. Front Plant Sci. 2024;15:1378421. 10.3389/fpls.2024.1378421.
16. Desta B, Getachew A. Paclobutrazol as a Plant Growth Regulator. Chem Biol Technol Agric. 2021;8(1):1–15. https://doi.org/10.1186/s40538-020-00199-z.
17. Dewi HSES, Prapto Y, Eka TSP, Benito HP. Physiological, Biochemical Activities of Cherelle Wilt on Three Cocoa Clones (*Theobroma cacao*) under Two Levels of Soil Fertilities. Biodiversitas. 2020;21:187–94.
18. Dewi ES. (2021) Tingkat Layu Pentil (*Cherelle Wilt*) pada Berbagai Klon Kakao (*Theobroma cacao* L). Jurnal Agropet 18(2).
19. Direktorat Jenderal Perkebunan. (2019). Statistik perkebunan Indonesia, outlook tanaman kakao 2019. Retrieved from https://ditjenbun.pertanian.go.id/template/uploads/2021/04/BUKU-STATISTIK-PERKEBUNAN-2019-2021-OK.pdf
20. Fan M, Moss KA, Jindal P, Kasznicki P, Davey P, Laissue PP, Lawson T. Large-scale quantification of stomatal patterning in barley leaves overexpressing epidermal patterning factor 1 reveals differential stomatal number between the adaxial and abaxial surfaces and spatial heterogeneity that impact stomatal function. New Phytol. 2025;248(4):1769–83. https://doi.org/10.1111/nph.70514.
21. Fuente De La GN, Frei UK, Trampe B, Ren J, Bohn M, Yana N, Verzegnazzi A, Murray SC, Lübberstedt T. (2019) A diallel analysis of a maize donor population response to in vivo maternal haploid induction II: Haploid male fertility. *Crop Sci* in press.
22. Ge Y, Hu X, Li X, Zhang W, Zhang C, Wu Z, Bai C, Ping Deng. Deep root in relation to grain yield and plant height of waxy maize. Chil J Agricultural Res. 2025;85(4). https://doi.org/10.4067/S0718-58392025000400565.
23. Guo Shuwei Z, Jiang N, Zhang S, Chen X, Qi Z, Ni C, Liu. One-step strategy for developing wheat haploid inducer lines with efficient markers for haploid identification. Crop J. 2026;2214–5141. https://doi.org/10.1016/j.cj.2026.03.010.
24. Ibrahim MM. Influence of Corms Size and Spraying with Benzyladenine and Paclobutrazol on Growth and Flowering Characteristics of *Freesia* sp. L. Plants. Swed J Sci Res. 2014;1(6):1–10. https://doi.org/10.13140/RG.2.1.3127.1526.

25. Indu Meena V, Saroj R, Patel M, Sharma D, Chand S, Chaudhary R, Singhal R, Rani R, Dadheech A. (2023) Doubled-Haploid Technology in Maize (*Zea mays* L.) and Its Practical Implications in Modern Agriculture. https://doi.org/10.1007/978-981-19-8218-7_6
26. Kamran M, Wennan S, Ahmad I, Xiangping M, Wenwen C, Xudong Z, Tiening L. Application of Paclobutrazol Affect Maize Grain Yield by Regulating Root Morphological and Physiological Characteristics Under a Semi-Arid Region. Sci Rep. 2018;8:1–5.
27. Kementerian Pertanian. *Outlook Komoditas Perkebunan Kakao 2022.* Pusat Data dan Sistem Informasi Pertanian Sekretariat Jenderal-Kementerian Pertanian. Indonesia: Jakarta; 2022.
28. Khammona K, Dermail A, Suriharn K, Lübberstedt T, Wanchana S, Thunnom B, Poncheewin W, Toojinda T, Ruanjaichon V, Arikit S. Accelerating haploid induction rate and haploid validation through marker-assisted selection for qhir1 and qhir8 in maize. Front Plant Sci. 2024;15:1337463. https://doi.org/10.3389/fpls.2024.1337463.
29. Kumari S, Bakshi P, Sharma A, Wali VK, Jasrotia A, Kour S. Use of Plant Growth Regulators for Improving Fruit Production in Sub Tropical Crops. Int J Curr Microbiol App Sci. 2018;7(3):659–68.
30. Li Yaoyao, Bo M, Liu PFY, Wang K, Hou P, Xue J, Li S, Xie R. Quantifying contributions of leaf area and longevity to leaf area duration under increased planting density and nitrogen input regimens during maize yield improvement. Field Crops Res Volume. 2022;283:0378–4290. https://doi.org/10.1016/j.fcr.2022.108551.
31. Lopez LA, Ochieng J, Pacheco M, Martinez L, Omar HA, Gowda M, Prasanna BM, Dhugga KS, Chaikam V. Effectiveness of R1-nj Anthocyanin Marker in the Identification of In Vivo Induced Maize Haploid Embryos. Plants (Basel Switzerland). 2023;12(12):2314. https://doi.org/10.3390/plants12122314.
32. Ma'arif S, Bambang SP, Arya WR, Iswati SD. Effects of shade on agronomic traits of doubled haploid lines of eggplant (*Solanum melongena* L.) obtained from anther culture. AIMS Agric Food. 2025;10(1):199–217. https://doi.org/10.3934/agrfood.2025011.
33. Magsayo NJ, Jamio E, Canunayon M, Tangpos D, Caritan J, Pascual PR. Root length and diameter at flag leaf stage correlate with important yield parameters in corn (*Zea mays* L.) grown in alkanline soil under drought conditions. ASEAN J Sci Tech Rep. 2025;28(5):e259419. https://doi.org/10.55164/ajstr.v28i5.259643.
34. Mandal P, Babu C, Natesan S, Kalichamy C, Deepika V, Hossain F, Ravikesavan R. From genes to table: genetic improvement of sweet corn for food quality and resilience. Euphytica. 2026;222. 10.1007/s10681-026-03688-8.
35. Marcondes MM, Schuck ML, Rossi RM. Effects of donor and inducer genotype and endosperm texture on haploid induction in maize assessed using R1-Navajo. Euphytica. 2026;222:81. https://doi.org/10.1007/s10681-026-03733-6.
36. Market Growth Report. (2026). Grain Market Size, Share, Growth, and Industry Analysis, By Type (Wheat, Rice, Corn, Barley), By Application (Food & Beverage, Animal Feed, Biofuel Production), Regional Insights and Forecast to 2035. India.

37. Martins T, Barros AN, Rosa E, Antunes L. Enhancing Health Benefits through Chlorophylls and Chlorophyll-Rich Agro-Food: A Comprehensive Review. Molecules. 2023;28(14):5344. https://doi.org/10.3390/molecules28145344.
38. Martínez-Barradas V, Francisco Barco-Rubio René, Morales-Poblete, Valentina V, Kattia N, Claudio I-B, Patricio, Arce-Johnson. Molecular and physiological regulation of stomatal function in legumes under abiotic stress: An integrative overview. Plant Stress. 2026;21:101348. https://doi.org/10.1016/j.stress.2026.101348.
39. Maqbool MA, Beshir A, Khokhar ES. Doubled haploids in maize: Development, deployment, and challenges. Crop Sci. 2020;60:2815–40. https://doi.org/10.1002/csc2.20261.
40. Messele MT, Temesgen Z, Solomon AD, Denekew BB, Hussein S. Linear mixed model to identify the relationship between grain yield and other yield related traits and genotype selection for sorghum. Heliyon. 2023;9(7):2405–8440. https://doi.org/10.1016/j.heliyon.2023.e17825. e17825.
41. Molenaar WS, Schipprack W, Brauner PC, Melchinger AE. Haploid male fertility and spontaneous chromosome doubling evaluated in a diallel and recurrent selection experiment in maize. Theor Appl Genet. 2019;132:2273–84.
42. Muenchrath D, Campbell A, Merrick L, Lübberstedt T, Fei S, Suza W. (2023) Ploidy: Polyploidy, aneuploidy, haploidy. In WP Suza, & KR Lamkey, editors, Crop Genetics. Iowa State University Digital Press. https://doi.org/10.31274/isudp.2023.130
43. Mushtaq A, Abida J, Monisa Y, Mudasir AM, Taha M, Tawheed A, Shabnum S, Asim R. Post-harvest quality management of sweet corn: Disorders, losses and preservation strategies. Food Nutr. 2025;1(1):100007. https://doi.org/10.1016/j.fnutr.2025.100007.
44. Nasaruddin, Tahir N, Ridwan I.) Fisiologi tumbuhan: Fitokhrom dan hormon pertumbuhan. Indonesia, Ficus Press: Makassar; 2019.
45. Binh-Anh NT, Thanh-Hao CL, Michael N, William RB, C. Engineering stomata for enhanced carbon capture and water-use efficiency. Trends Plant Sci. 2023;28(11):1290–309. https://doi.org/10.1016/j.tplants.2023.06.002. ISSN 1360–1385.
46. Oktaviani W. (2008) Peningkatan Produksi Buah Kakao (*Theobroma cacao* L.) melalui Pemberian Zat Pengatur Tumbuh Paclobutrazol pada berbagai Konsentrasi. Bachelor's thesis. Retrieved from http://repository.ipb.ac.id/handle/123456789/2991
47. Oyekale AS. Determinants of Cocoa Farmers Compliance with Agrochemical Safety Precautions in Ogun and Osun States, Nigeria. Toxics. 2022;10(8):454. https://doi.org/10.3390/toxics10080454.
48. Pan S, Xin W, Zhengbing Y, Jin W, Lulu G, Ziyang P, Yuntao W, Jing L, Bin W, Yanjun S, Lingli L. Leaf stomatal configuration and photosynthetic traits jointly affect leaf water use efficiency in forests along climate gradients. New Phytol. 2024;244(4):1250–62. https://doi.org/10.1111/nph.20100.
49. Putro SAC, Yusalina, Ratna W. (2016) Factors Affecting Export of Indonesia Cocoa Beans to Malaysia. Makassar, Indonesia. Jurnal Manajemen & Agribisnis 20(1): 90–101.
50. Rahman M, Noor AS, Nurniwalis AW, Puteri EMW, Abdullah MO, Nur APA, Ghulam KAP, Jeremy AR, Zubaidah R. Impact of Paclobutrazol on The Growth and Development of Nursery Grown Clonal Oil

Palm (*Elaeis guineensis* Jacq). J Oil Palm Res. 2016;28:404–14. 10.21894/jopr.2016.2804.02.

51. Ribeiro MR, Roberto dosST, Isabel RP, de S, Andréa AC, Tácila Cde, Azevedo, Lauro JMG, Saulo F, da SC, Luiz A, dos SD. Are stomatal area and stomatal number reliable traits for identification of doubled haploids in maize? Crop Breed Appl Biotechnol. 2022;22(2). https://doi.org/10.1590/1984-70332022v22n2a16.
52. Rozi F, Santoso AB, Mahendri IGAP, Hutapea RTP, Wamaer D, Siagian V, Elisabeth DAA, Sugiono S, Handoko H, Subagio H, Syam A. Indonesian market demand patterns for food commodity sources of carbohydrates in facing the global food crisis. Heliyon. 2023;9(6):e16809. https://doi.org/10.1016/j.heliyon.2023.e16809.
53. Ruanjaichon V, Khammona K, Thunnom B, Suriharn K, Kerdsri C, Aesomnuk W, Yongsuwan A, Chaomueang N, Thammapichai P, Arikit S, Wanchana S, Toojinda T. Identification of Gene Associated with Sweetness in Corn (*Zea mays* L.) by Genome-Wide Association Study (GWAS) and Development of a Functional SNP Marker for Predicting Sweet Corn. Plants. 2021;10(6):1239. https://doi.org/10.3390/plants10061239.
54. Rünneburger E, Le Rouzic A. Why and how genetic canalization evolves in gene regulatory networks. BMC Evol Biol. 2016;16:239. https://doi.org/10.1186/s12862-016-0801-2.
55. Sabooni N, Gharaghani A. Induced polyploidy deeply influences reproductive life cycles, related phytochemical features, and phytohormonal activities in blackberry species. Front Plant Sci. 2022;13:938284. https://doi.org/10.3389/fpls.2022.938284.
56. Sekiya A, Pestana JK, Silva MGB, Da Krause MD, Silva CRM, Da Ferreira JM. Haploid induction in tropical supersweet corn and ploidy determination at the seedling stage. Pesquisa Agropecuária Brasileira. 2020;55:e00968. https://doi.org/10.1590/S1678-3921.pab2020.v55.00968.
57. Sherif HM, Asaad SA. Effect of Some Plant Growth Retardants on Vegetative Growth, Spurs and Fruiting of 'Le-Cnte' Pear Tress. Br J Appl Sci Technol. 2014;4(26):3785–804.
58. Shi X, Hua Y, Chen C, Jie H, Katherine MH, Patrice SA, Tieming J, Jianlin C, James AB. Genomic imbalance determines positive and negative modulation of gene expression in diploid maize. Plant Cell. 2021;33(4):917–39. https://doi.org/10.1093/plcell/koab030.
59. Sidahmed H, Vad A, Nagy J. Advances in Sweet Corn (*Zea mays* L. saccharata) Research from 2010 to 2025: Genetics, Agronomy, and Sustainable Production. Agronomy. 2025;15(5):1260. https://doi.org/10.3390/agronomy15051260.
60. Silva HAda, Carlos AS, Janieli MSV, Antonio TA, Júnior, Ronald JBP, Káthia SMM, Robson MR, Alessandra GB. Effect of ploidy level on guard cell length and use of stomata to discard diploids among putative haploids in maize. Crop Sci. 2019;60(3):1199–209. https://doi.org/10.1002/csc2.20004.
61. Sinay H. (2018) Pengaruh Giberelin dan Temperatur terhadap Pertumbuhan Semai Gandaria (*Bouea macrophylla* Griffth). Bioscientiae 8(1).
62. Sitohang N, Erwin MH, Chairani H, Tumpal HSS, Hasril S. Responses such as Flushing, Flowering, Pod Reserving, and Yield of TSH 858 Clone Cacao (*Theobroma cacao* L.) to an Increase in Dose of

Balanced N.P.K.Ca.Mg 12.9:11.4:16.8:10.6:4.8 Fertilizing. J Hunan Univ (Natural Sciences). 2021;48(2):165–74.

63. de Souza LdeF, Bruno FF, Juliana, Moraes M, de O, Vitor J, de Otavio L, Iran de GLH, Josué AD. MF (2024) Early haploid identification by stomatal guard cell length in tropical supersweet corn using different inducers. Crop Breeding and Applied Biotechnology, 24(1): e46052414. Brazilian Society of Plant Breeding. http://dx.doi.org/10.1590/1984-70332024v24n1a04
64. Subaedah St, Edy E, Mariana K. Growth, Yield, and Sugar Content of Different Varieties of Sweet Corn and Harvest Time. Int J Agron. 2021;8882140:7. https://doi.org/10.1155/2021/8882140.
65. Tandel YN, Patel NL. Effect of Chemicals on Growth, Yield and Economics of Mango (*Mangifera indica* L). Kartanaka Agric Sci. 2011;24(3):362–5.
66. Taskin SZ, Ugur B. Development of Maize (*Zea mays* L.) Genotypes by using In Vivo Doubled Haploid Technique. Turk J Field Crops. 2023;28(1):1–6. https://doi.org/10.17557/tjfc.1218958.
67. Uliana TH, Batîru G, Frei UK, Dutta S, Lübberstedt T. Investigating the Effect of the Interaction of Maize Inducer and Donor Backgrounds on Haploid Induction Rates. Plants (Basel Switzerland). 2022;11(12):1527. https://doi.org/10.3390/plants11121527.
68. Ulyanov AV, Karlov AV, Khatefov EB. The use of maize haploidy inducers as a tool in agricultural plant biotechnology. Vavilovskii zhurnal genetiki i selektsii. 2022;26(7):704–13. https://doi.org/10.18699/VJGB-22-85.
69. Vanous K, Talukder ZJ, Ursula KF, Baskar G, Thomas L. Utilization of Reduced Haploid Vigor for Phenomic Discrimination of Haploid and Diploid Maize Seedlings. Plant Phenome J. 2019;2(1):180008. https://doi.org/10.2135/tppj2018.10.0008.
70. van Vliet JA, Giller KE. Chapter Five - Mineral Nutrition of Cocoa: A Review. Adv Agron. 2017;141:185–270. https://doi.org/10.1016/bs.agron.2016.10.017.
71. Waters AJ, Bilinski P, Eichten SR, Vaughn MW, Ross-Ibarra J, Gehring M, Springer NM. Comprehensive analysis of imprinted genes in maize reveals allelic variation for imprinting and limited conservation with other species. Proc Natl Acad Sci USA. 2013;110(48):19639–44. https://doi.org/10.1073/pnas.1309182110.
72. Xia X, Yuhan T, Wei M, Daqiu Z. Effect of Paclobutrazol Application on Plant Photosynthetic Performance and Leaf Greenness of Herbaceous Peony. Horticulturae. 2018;4(1):5. https://doi.org/10.3390/horticulturae4010005.
73. Xue R, Zuo W, Zheng Z, Han Q, Shi J, Zhang Y, Qiu J, Wang S, Zhu Y, Cao W, Zhang X. Interpreting Controls of Stomatal Conductance across Different Vegetation Types via Machine Learning. Water. 2024;16(16):2251. https://doi.org/10.3390/w16162251.
74. Yang Ji-wei, Zong-hua L, Yan-zhi Q, Ya-zhou Z, Hao-chuan L. Cytological study on haploid male fertility in maize. J Integr Agric. 2022;21(11):3158–68. https://doi.org/10.1016/j.jia.2022.07.055. ISSN 2095–3119.
75. Yavuz R, Wang H, Fakude M, Dermail A, Frei UK, Liu P, Lübberstedt T. Genetic variation and heritability of haploid frailty in maize. Front Plant Sci. 2025;16:1572901.

https://doi.org/10.3389/fpls.2025.1572901.

76. Zhao L, Luo LT, Yue EK, Shi J. Integrating haploid breeding and germplasm innovation in maize disease resistance breeding, case study. Mol Plant Breed. 2025;16(2):146–55. http://doi.org/10.5376/mpb.2025.16.0015.
77. Zhu Y, Junhao W, Yongjun Z, Han L, Ziren Q, Jian Z, Dan W, Pu W, Qingfeng M. Unlocking maize yield potential through exploring canopy-root interactions with nature-based nutrient management. Field Crops Res. 2024;318(109618):0378–4290. https://doi.org/10.1016/j.fcr.2024.109618.

Figures

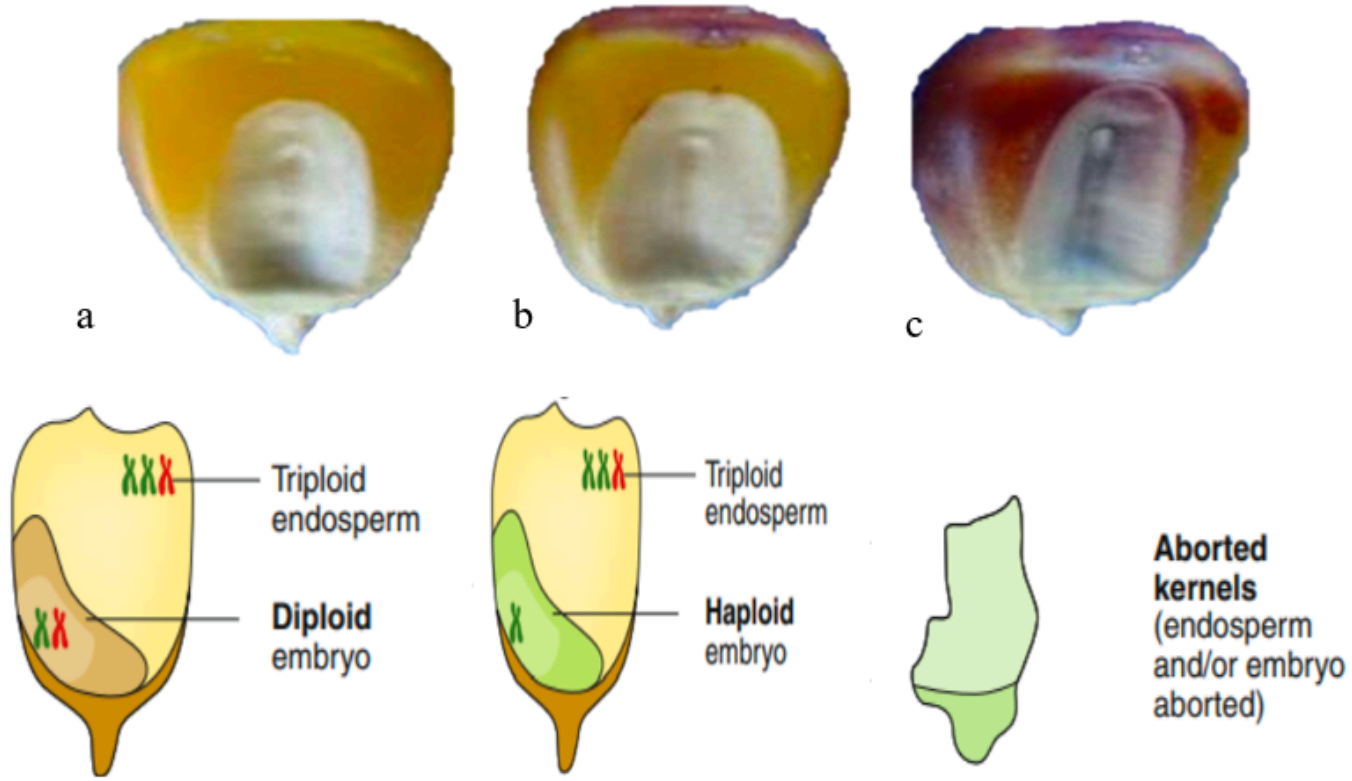


Figure 1

Seed Morphological (a) Inbreed; (b) Haploid; (c) Hybrid

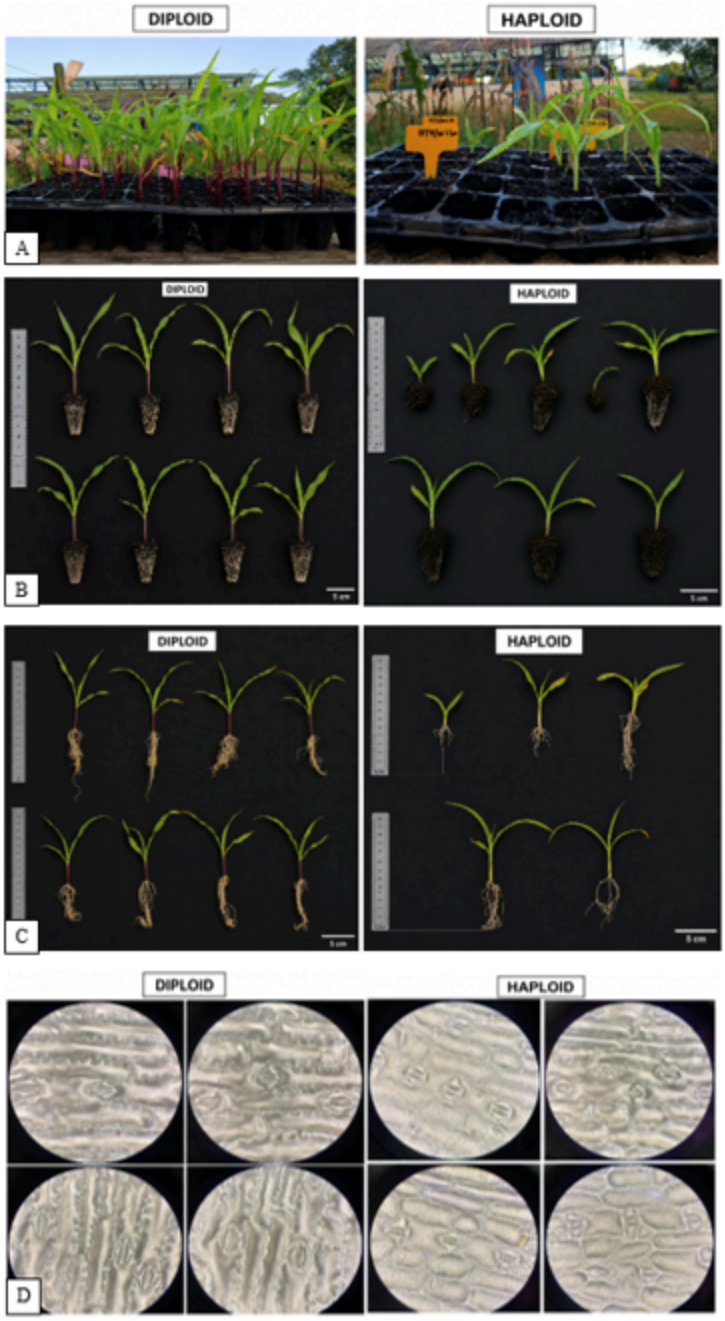


Figure 2

Representative phenotypic and anatomical characteristics of diploid and haploid sweet corn seedlings. (A) Whole-seedling morphology showing differences in plant vigor and overall growth. (B) Representative leaf morphology illustrating variation in leaf development. (C) Root system architecture showing differences in root development and lateral root formation. (D) Representative microscopic images of stomatal characteristics.

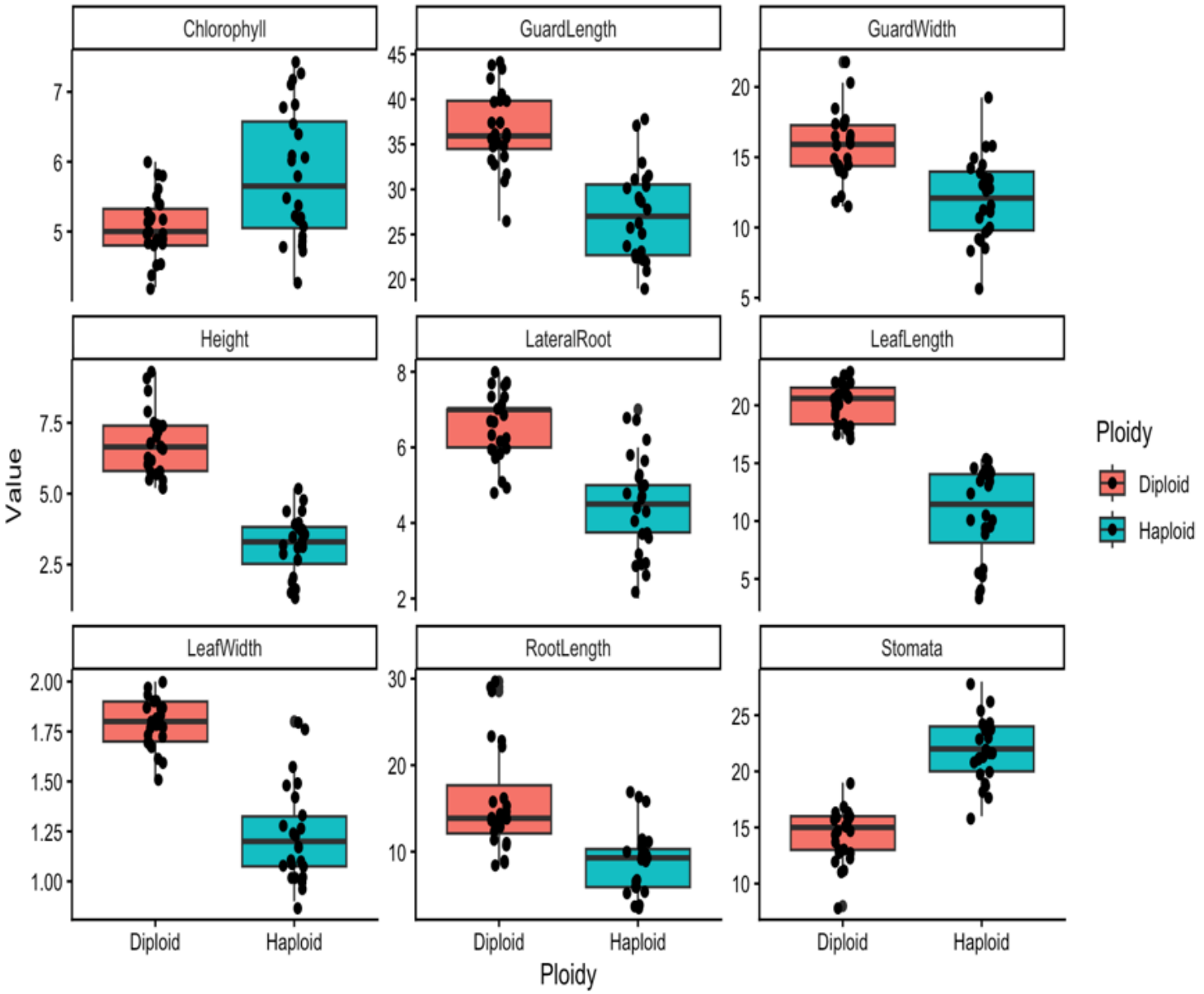


Figure 3

Distribution of morphological and physiological traits between haploid and diploid sweet corn seedlings.

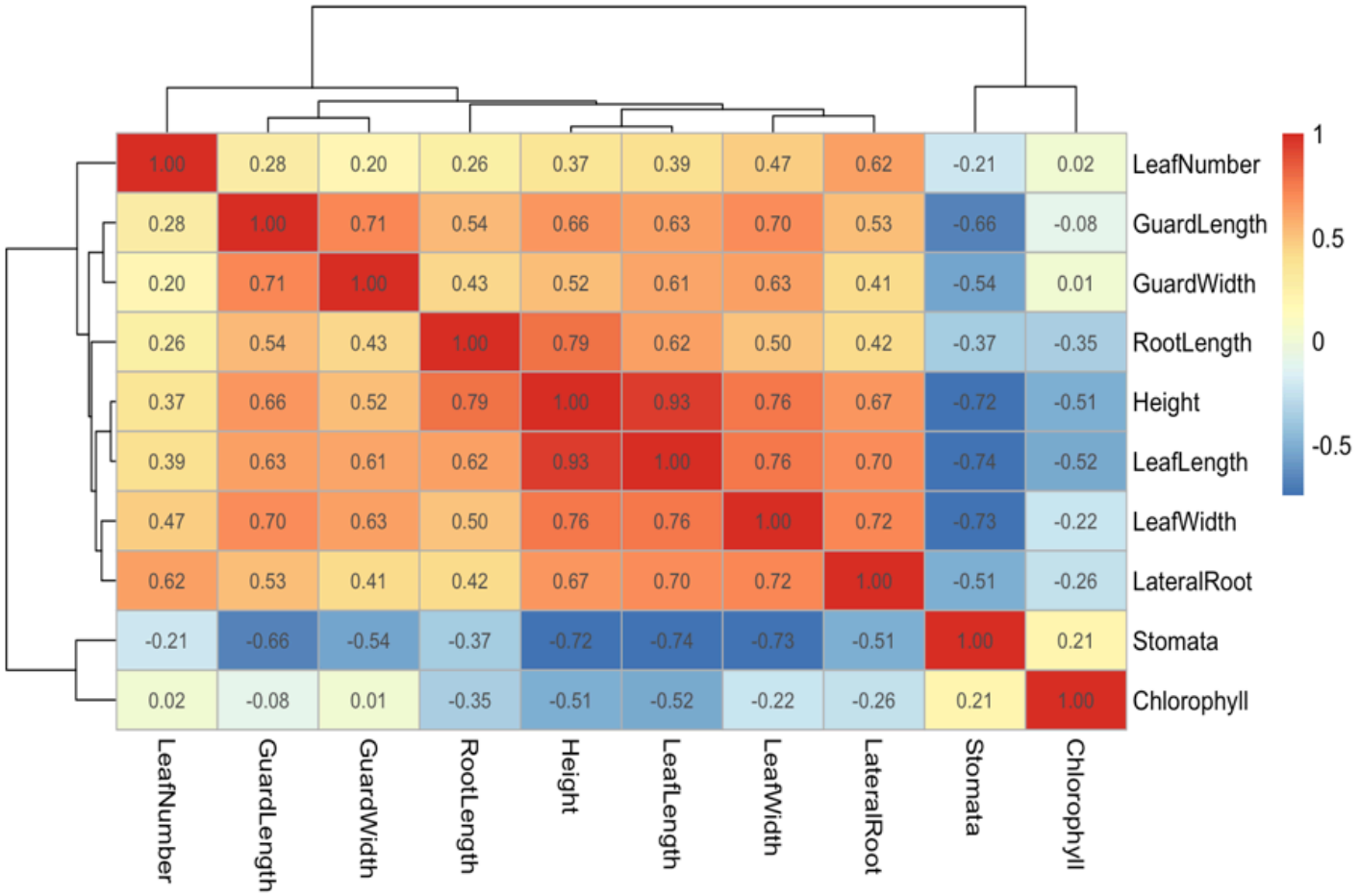


Figure 4

Pearson correlation matrix of morphological and physiological traits in sweet corn seedlings.

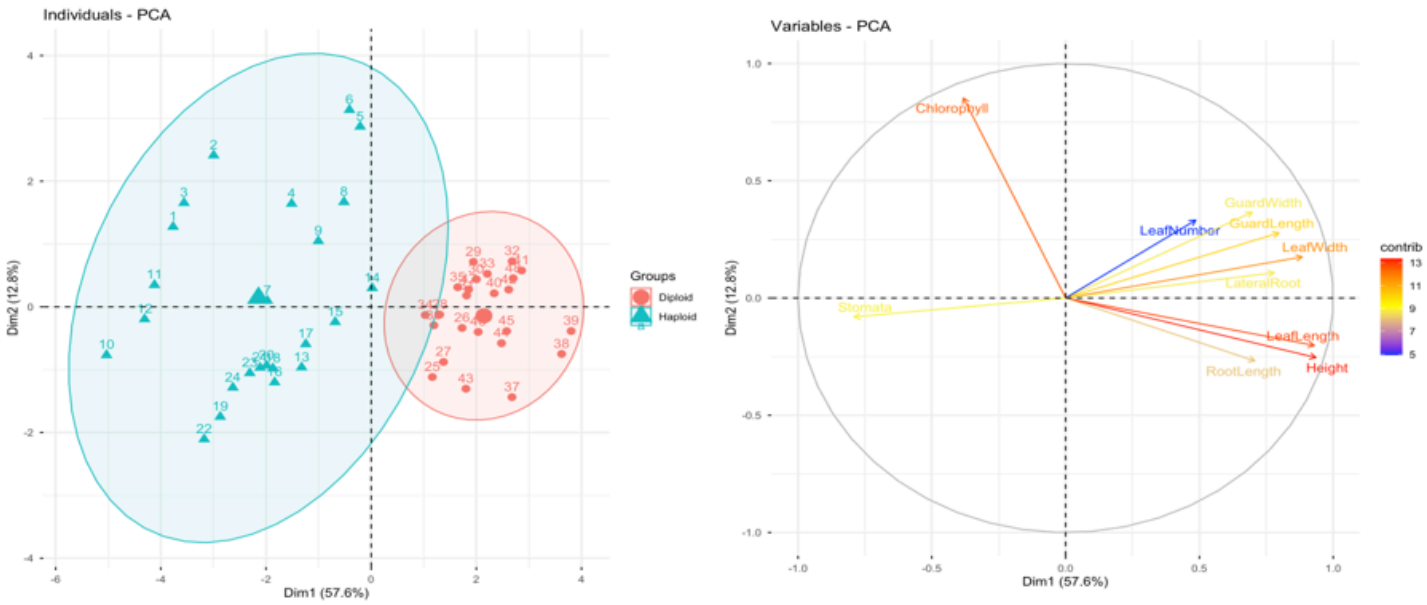


Figure 5

(a) Principal component analysis of sweet corn seedlings according to routine haploid and diploid classification (b) Contribution of phenotyphic traits to the first two principal components.

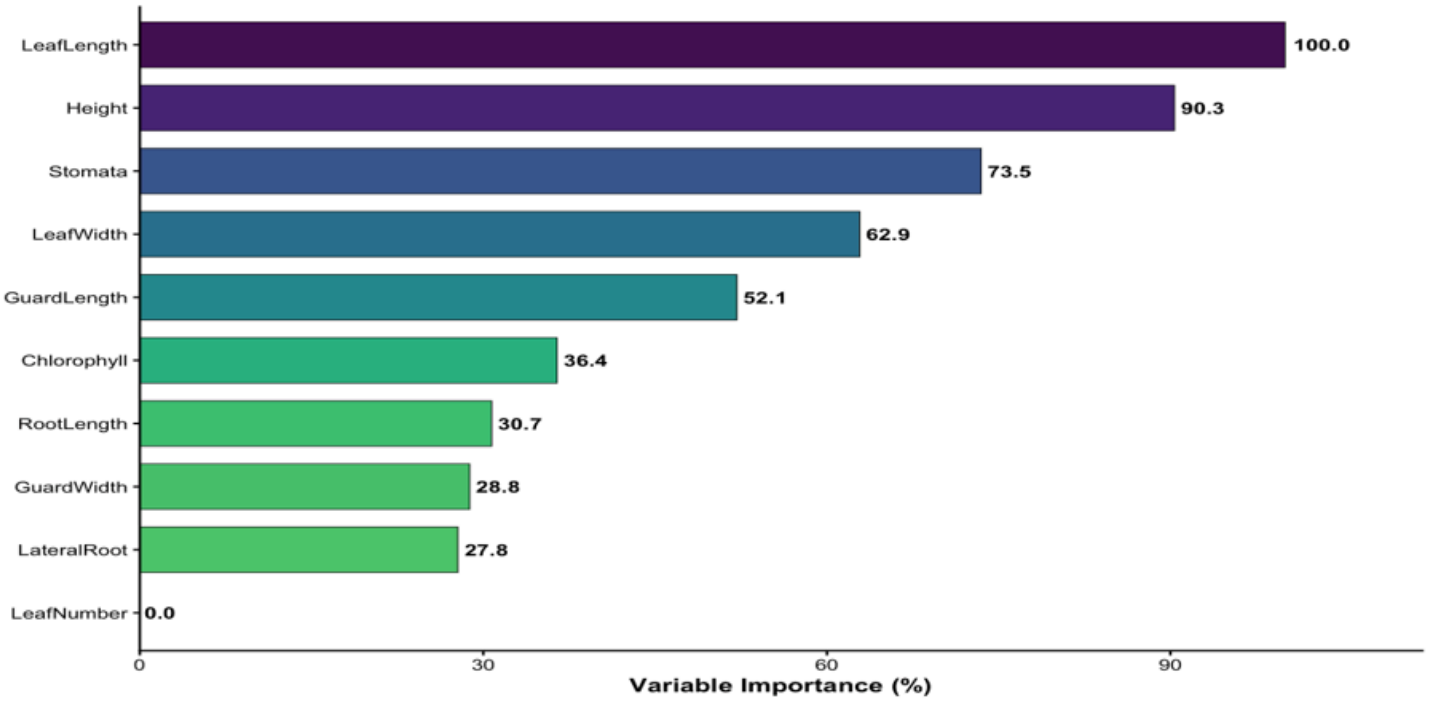


Figure 6

Random forest variable importance for discrimination of seedlings according to routine haploid and diploid classification.

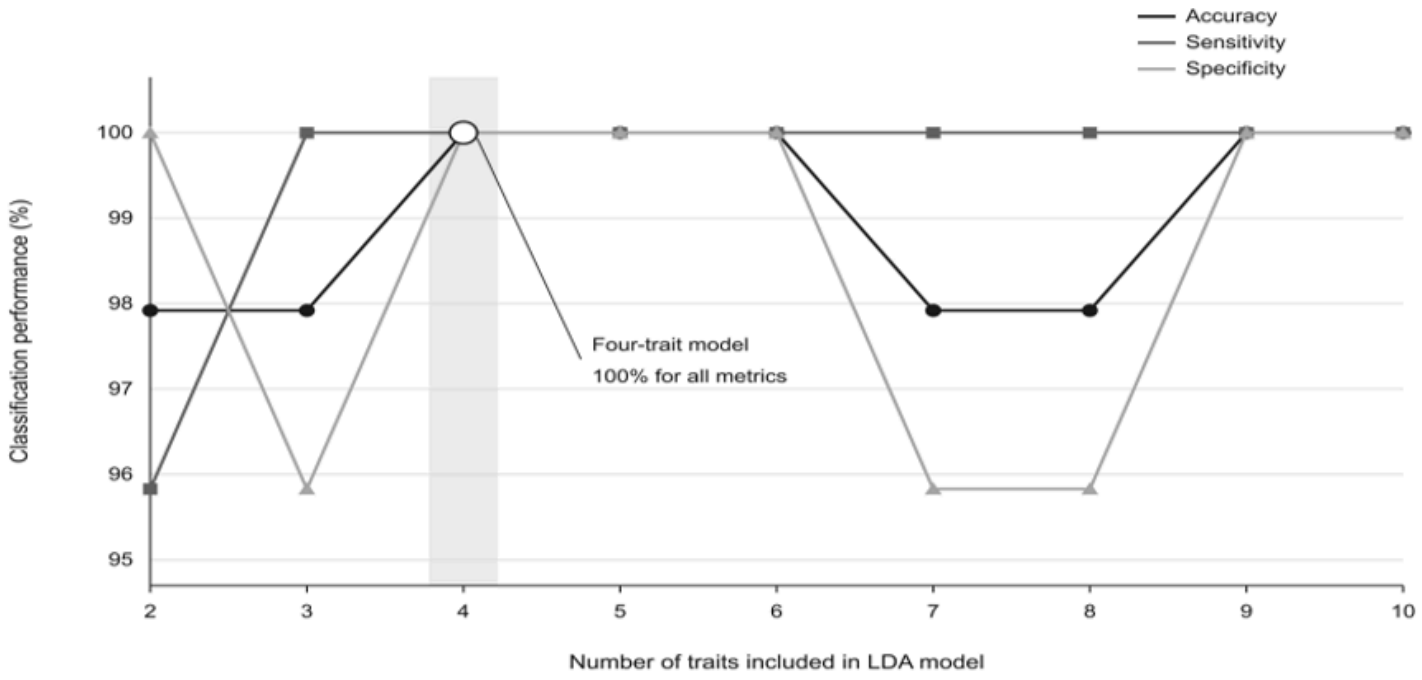


Figure 7

Classification performance of LDA models constructed with 2–10 phenotypic traits, added sequentially according to Random Forest importance.

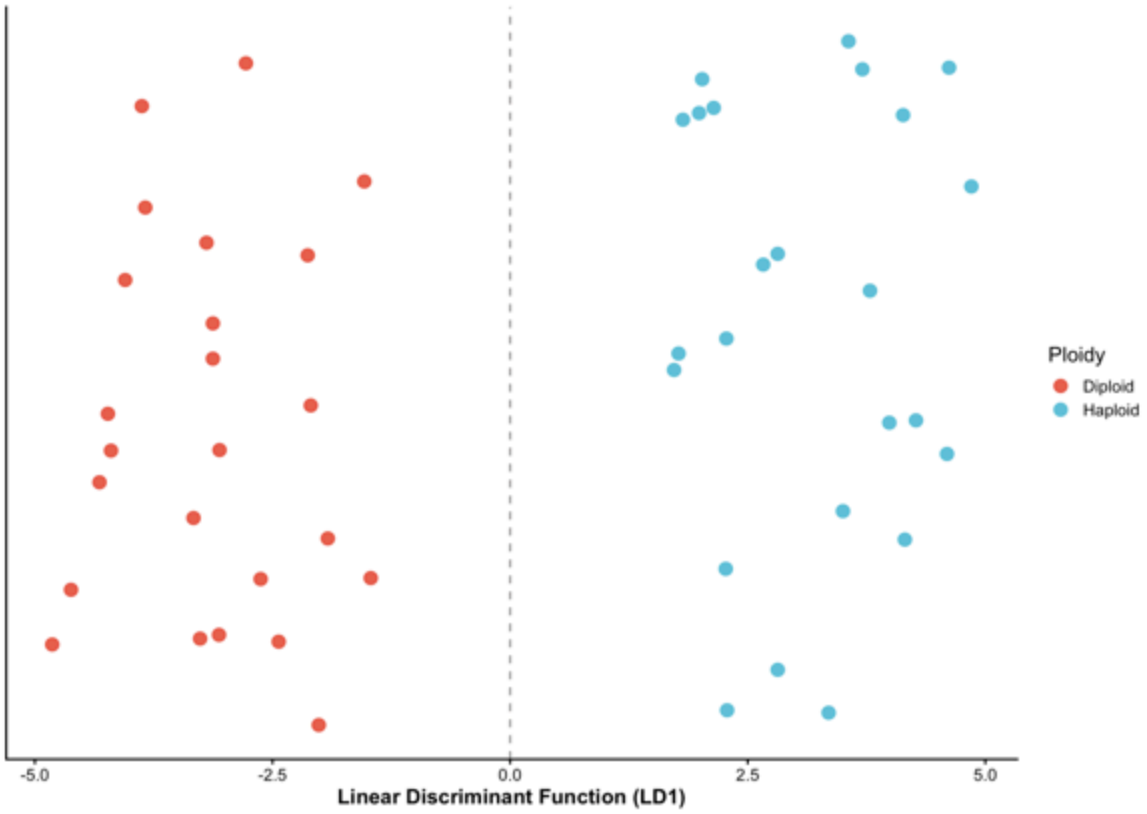


Figure 8

Linear discriminant analysis (LDA) showing phenotypic separation of seedlings according to routine haploid and diploid classification.