in silico Plants, 2025, 7, diaf017
https://doi.org/10.1093/insilicoplants/diaf017
Advance access publication 7 October 2025
Original Research

Original Research

A robust model based on root morphological and anatomical features to distinguish high and low methane emission rice varieties through machine learning approaches

Ripon Kumar Roy[1,2,3,‡], Seyed Mahdi Hosseiniyan Khatibi[1,‡], Sung-Ryul Kim[1], Kurniawan Rudi Trijatmiko[1], Maria Genaleen Q. Diaz[4], Eureka Teresa M. Ocampo[3], Jose E. Hernandez[3], Saeed Pirmoradi[5], Amelia Henry[1,*], Ajay Kohli[1,*]

[1]Rice Breeding Innovations Department, International Rice Research Institute (IRRI), Pili Drive, UPLB Campus, Los Baños, Laguna 4031, Philippines
[2]Biotechnology Division, Bangladesh Rice Research Institute (BRRI), Gazipur-1701, Bangladesh
[3]Institute of Crop Science, College of Agriculture and Food Science, University of the Philippines Los Baños, College, Laguna 4031, Philippines
[4]Institute of Biological Sciences, College of Arts and Sciences, University of the Philippines Los Baños, College, Laguna 4031, Philippines
[5]Department of Computer Science, Khajeh Nasir Toosi University of Technology, 470 Mirdamad Avenue West, Tehran 19697-64499, Iran

[‡]These authors contributed equally and should be considered as co-first authors.
[*]Corresponding authors. Amelia Henry, Rice Breeding Innovations Department, International Rice Research Institute, Pili Drive, UPLB Campus, Los Baños, Laguna 4031, Philippines.
E-mail: a.henry@cgiar.org; Ajay Kohli, Rice Breeding Innovations Department, International Rice Research Institute, Pili Drive, UPLB Campus, Los Baños, Laguna 4031, Philippines.
E-mail: ajakoy@gmail.com

Handling Editor: Dr. Xin-Guang Zhu

Abstract. Rice fields are a major producer of methane, a strong greenhouse gas. However, identifying genetic variation in methane emissions among rice varieties remains challenging. This study applied association rule mining to detect key rice root morphological and anatomical traits influencing methane emissions, validated using a support vector machine. We report models which accurately classified high and low methane-emitting varieties with 98% (morphological) and 94% (anatomical) accuracy. These models effectively distinguished methane emission categories based on intrinsic trait patterns. Machine learning analysis highlighted the top 10 morphological and anatomical traits associated with methane emission levels. High methane-emitting varieties were characterized by lower middle root porosity, base root porosity, average root porosity, root diameter (RDia), and higher S-type lateral root length. Conversely, low methane-emitting varieties exhibited lower root number, tiller number, root dry weight, leaf number, and higher RDia. Anatomically, high methane-emitting varieties showed reduced lacunae number, total stele area, mean metaxylem size, metaxylem number, and metaxylem vessel area. Low methane-emitting varieties, in contrast, had higher percent aerenchyma, total stele area, ratio of total cortical area to root cross-section area, ratio of stele to root cross-section area, and aerenchyma area. The results suggest that the rhizosphere oxygenation role of root porosity and aerenchyma might predominate over the methane transport role.

Keywords: methane emission; high methane-emitting rice variety; low methane-emitting rice variety; morphological traits; anatomical traits; machine learning; model

1. INTRODUCTION

Rice (*Oryza sativa* L.) farming is widely acknowledged as a major source of methane, a potent greenhouse gas. Estimates of methane output from rice paddies worldwide range between 493 and 723 million tons of CO_2-equivalent annually (Edenhofer et al. 2014), representing about 30% of global agricultural methane emissions (Gupta et al. 2021). Rice plants are thought to significantly contribute to this process (Das and Baruah 2008a, 2008b, Qin et al. 2015, Su et al. 2015) primarily due to their well-developed aerenchyma; specialized gas-conducting tissues found in the leaves, stems, and roots that facilitate the internal transport of gases within the plant (Colmer 2003b). A significant share of methane emissions is thought to occur through the plants via aerenchyma in rice roots, with

60%–98% of methane reported to be emitted through the aerenchyma route into the atmosphere (Denier van der Gon and Neue 1996, Butterbach-Bahl et al. 1997b, Wassmann et al. 2000, Philippot et al. 2011, Wang et al. 2017). Methane emission levels have been reported to differ among rice cultivars, which might be a key consideration towards emission reduction (Qin et al. 2015, Su et al. 2015); across studies in different locations, some varieties consistently released lower levels of methane compared to others grown under similar conditions (Butterbach-Bahl et al. 1997a, Shin and Yun 2000, Singh et al. 2003, Lou et al. 2008, Baruah et al. 2010, Ghani Nugroho et al. 1997). Therefore, classification of rice varieties as low or high methane-emitting was possible.

A number of plant characteristics specific to the rice variety may be associated with methane emissions. Morphological (Ma et al. 2010), physiological (Jiang et al. 2019), and anatomical (Das and Baruah 2008c) features have all been proposed as factors influencing methane emissions from rice fields, although findings across studies have not always been consistent. For example, rice varieties with smaller xylem vessels were identified as low methane-emitting (Bharali et al. 2017b). Traits such as leaf number, leaf area, root volume, root length, root number, and tiller number have shown positive correlations with methane emissions (Das and Baruah 2008a), along with total plant biomass (Sass et al. 1990, Cao et al. 1996, Huang et al. 1997). Moreover, hybrid rice varieties have been reported to produce 50%–60% more above-ground biomass than traditional inbred varieties, which may enhance oxygen supply in the root zone and thereby promote methane oxidation in the soil (Ma et al. 2010). On the other hand, some studies have observed that certain traits can contribute to decreasing methane emissions, including increased root length, root biomass, root dry weight, as well as enhanced root oxidation capacity, and radial oxygen loss (Chen et al. 2019), along with greater total and active absorbing surfaces (Zhang et al. 2019). Photosynthesis is also thought to influence methane production in rice paddies, as carbon compounds generated during photosynthesis serve as substrates for methanogenic microorganisms in the root zone (Das and Baruah 2008a). Additionally, specific root exudates, such as malic acid, succinic acid, and citric acid, have been reported to support the growth and activity of methanotrophic bacteria, thereby aiding methane oxidation in the rhizosphere (Chen et al. 2019, Zhang et al. 2019).

Oxygen availability in the rhizosphere plays a crucial role in regulating methane production; as such, root traits associated with oxygen transport are likely important determinants of methane emissions in rice (Jiménez and Pedersen 2023). Radial oxygen loss (ROL) enhances aerobic conditions around the roots, and rice varieties characterized by a greater root diameter, biomass, number, and porosity tend to release more oxygen into the surrounding soil (Li and Wang 2013). A negative relationship between methane emissions and both root porosity and oxygen release capacity has been reported (Fang et al. 2021). Rice varieties that exhibit elevated ROL levels have been found to emit less methane (Duyen et al. 2022), likely due to improved oxygen diffusion in the rhizosphere, which inhibits methanogenic activity (Duyen et al. 2022). Furthermore, ROL was positively linked with root traits including a higher number of roots per plant, increased root tips, and larger root volume (Zheng et al. 2018). Consistent findings by Chen et al. (2019) indicated that rice varieties having greater root dry weight, length, oxidation potential, and ROL tended to produce lower levels of methane. These root traits likely facilitate more efficient oxygen transfer into the soil, thereby promoting methane oxidation. Increased root biomass and porosity were reported to improve oxygen availability in the rhizosphere, contributing to reduced methane emissions (Li et al. 2022). More recently, Qi et al. (2024) reported that rice varieties with greater root porosity and oxygen release capacity emitted significantly less methane. In contrast, the formation and structural characteristics of aerenchyma in rice roots and shoots are critical in regulating the plant`s capacity to transport methane from the rhizosphere the atmosphere. Aulakh et al. (2000) reported that rice cultivars with more extensive aerenchyma networks exhibited significantly higher methane emissions, which was attributed to enhanced internal gas transport pathways. Supporting this, Aulakh et al. (2001a, 2001b) found that cultivars with larger aerenchyma lacunae possessed greater methane transport capacities, emphasizing the influence of anatomical features on gas exchange efficiency. Furthermore, aerenchyma formation is reportedly modulated by soil redox potential, with reduced (more anaerobic) conditions stimulating increased development of these tissues, thereby facilitating greater methane release from soil to the atmosphere (Kludze et al. 1993). In addition to anatomical traits, morphological characteristics also impact methane dynamics; Das and Baruah (2008b) reported that rice varieties with greater leaf area and tiller numbers—traits associated with higher aerenchyma density—tended to emit more methane, highlighting the integrated role of plant structure in methane emission processes.

These contrasting reports highlight how, despite the long list of potential traits responsible for rice methane emissions, there are distinct controversies in defining the traits responsible for genetic variation resulting in the classification of rice varieties as high and low methane-emitting. This lack of clarity may be in part due to the design of previous methane studies, for which the test varieties were selected without knowing whether they were high or low methane-emitting, and a relatively high methane-emitting variety in one study might be considered a low methane-emitting variety in another depending on the other entries in the study. Furthermore, differences in site characteristics and interactions among traits may affect methane emissions results among studies. Previous studies characterizing rice genetic variation in methane emissions have included comparisons of traditional and improved varieties (Aulakh et al. 2000, Das and Baruah 2008a, 2008c) wetland species (Lai et al. 2012), high-yielding summer rice varieties (Bharali et al. 2017b), high yielding and low yielding rice varieties (Jiang et al. 2017, Bharali et al. 2017a) improved inbred varieties (Huang et al. 1997), and *indica*, *japonica*, and hybrid rice varieties (Ma et al. 2010, Chen et al. 2019).

An experimental design aimed specifically to contrast the traits exhibited by high and low methane-emitting varieties may result in a more clear distinction between the two varietal groups. In our study, we used 14 varieties that had been reported as high methane-emitting and 11 varieties that had been reported as low methane-emitting in previous studies (Table 1). Given the lack of consensus in the literature regarding which plant traits

Table 1. List of varieties used in this study that were classified as high and low methane emitting.

SL no.	Name	IRGC no.	Source country	Status of methane emission	References
1	Monohar sali	47107	India	High	Gogoi et al. (2008)
2	Ratna	78608	India	High	Adhya et al. (1994)
3	Pusa-33	74331	India	High	Singh et al. (1997)
4	Pusa-169	74332	India	High	Singh et al. (1997)
5	Mahsuri	52696	Malaysia	High	Baruah et al. (2010)
6	Della	117706	USA	High	Ding et al. (1999)
7	Maybelle	78629	USA	High	Lindau et al. (1995)
8	Tebonnet	66760	USA	High	Lindau et al. (1995)
9	Mars	38541	USA	High	Lindau et al. (1995), Ding et al. (1999)
10	Cisadane	69983	Indonesia	High	Setyanto et al. (2004)
11	Membramo	—	Indonesia	High	Setyanto et al. (2004)
12	Way Apo Buru	—	Indonesia	High	Setyanto et al. (2004)
13	Dular	10615	India	High	Wang et al. (1997)
14	Annada	120710	India	High	Adhya et al. (1994)
15	Lemont	117802	USA	Low	Ding et al. (1999)
16	Labelle	24274	USA	Low	Ding et al. (1999)
17	Dawn	2026	USA	Low	Ding et al. (1999)
18	Katy	126386	USA	Low	Ding et al. (1999)
19	Lebonnet	24275	USA	Low	Ding et al. (1999)
20	IR 36	30416	Philippines	Low	Ding et al. (1999), Gogoi et al. (2008)
21	Cypress	124359	USA	Low	Ding et al. (1999), Lindau et al. (1995)
22	Bengal	126402	USA	Low	Lindau et al. (1995)
23	Lacassine	122146	USA	Low	Lindau et al. (1995)
24	IR72	76330	Philippines	Low	Wang et al. (1997)
25	IR64	116793	Philippines	Low	Setyanto et al. (2004), Gorh and Baruah (2019)

most strongly influence methane emissions, our objectives were to (i) investigate the morphological and anatomical differences in high vs low methane-emitting varieties, (ii) introduce significant traits in high/low methane-emitting varieties based on morphological and anatomical traits, and (iii) design a robust model for predicting and classifying high/low methane-emitting rice varieties based on morphological and anatomical characteristics through machine learning approaches.

2. MATERIALS AND METHODS

2.1 Selection of high and low methane-emitting varieties

Rice varieties were selected based on information from published studies (Supplementary Table S1). All experiments had been conducted under continuous flooding conditions, typically beginning after plant establishment and maintaining 5–10 cm of standing water throughout the growing period. We prioritized sources that categorized varieties as high or low methane-emitting based on multi-seasonal comparisons—such as the studies by Ding et al. (1999) and Gorh and Baruah (2019)—or across different soil types, as reported by Singh et al. (1997) and Ding et al. (1999). Preference was also given to varieties consistently categorized across multiple sources. For example, the variety 'Mars' was identified as high methane emitting by both Lindau et al. (1995) and Ding et al. (1999). Similarly, 'IR36' was consistently reported as low methane emitting in the studies by Ding et al. (1999), Adhya et al. (1994), and Gogoi et al. (2008). 'Cypress' was also classified as low methane-emitting by both Ding et al. (1999) and Lindau et al. (1995). Although methane emission levels varied across years, experiments, and soil types, the overall trend remained consistent in the high and low methane-emitting varieties. Availability of the varieties in the IRRI Genebank was also a key consideration for selection.

2.2 Experiment setup and data preparation

Fourteen and eleven rice varieties that were previously classified as high and low methane-emitting, respectively (Table 1), were grown in a greenhouse study at the International Rice Research Institute (IRRI), Los Baños, Laguna, Philippines. The seeds were provided by the IRRI Genebank. Four kilograms of sandy loam soil containing 14% moisture was filled into 6-l green pots, with 2.5 g of complete fertilizer (14:14:14-NPK) mixed with the soil in each pot. The experiment was conducted in a completely randomized design (CRD), with each rice variety replicated three times (one plant per pot). Throughout the experiment, temperature and relative humidity were monitored using a HOBO Pro v2 U23-001A datalogger (Onset USA).

2.3 Morphophysiological traits

Twenty morphophysiological traits (Table 2) were measured on three replicate plants per variety. Non-destructive chlorophyll content index (CCI: CCM-200, Apogee Instruments), was measured at three different positions on each leaf. Stomatal conductance (CD) and photosynthesis measurements (portable photosynthesis system, LI-6400) were performed on the two youngest, healthy, and fully expanded leaves from each plant. Prior to harvest, tiller number (TN), and leaf number (LN) per plant were recorded.

Sampling of root and shoot tissue was based on the expected time of peak methane emissions. Two major peaks of methane

Table 2. Morphological and anatomical traits measured on 14 and 11 rice varieties previously classified as high and low methane-emitting, respectively.

SL no.	Morphophysiological traits	Anatomical traits
1	Base Root Porosity (**BP**)	Root Cross-section Area (**RXSA**-mm^2)
2	Middle root Porosity (**MP**)	Total Cortical Area (**TCA**-mm^2)
3	Tip Root Porosity (**TP**)	Total Stele area (**TSA**-mm^2)
4	Average Root Porosity (**RP**-Average of Base, Middle and Tip Root Porosity)	Ratio of Total Cortical Area to Stele Area (**RCS**)
5	Root Diameter (**Rdia**-mm)	Ratio of Total Cortical Area to Root Cross-Section Area (**RCXS**)
6	Root tips (**Tips**)	Ratio of Stele to Root Cross-Section Area (**RSXS**)
7	Total Root Length (**TRL**-cm m^{-3})	Aerenchyma Area (**AA**-mm^2)
8	Percent S-Type Lateral Root (**STLR**)	Percent Live cortical Cell (**PLCC**- Unbroken cortical cell)
9	Percent L-Type Lateral Root (**LTLR**)	Percent Aerenchyma (**PA**)
10	Percent Lateral Root (**PLR**)	Metaxylem Vessel Area (**MXVA**-mm^2)
11	Percent Nodal Root (**PNR**)	Percent Metaxylem Vessel Area in Stele Area (**PMXVAS**)
12	Leaf Number (**LN**)	Percent Metaxylem Vessel Area in Root Cross-Section Area (**PMXVA-XSA**)
13	Tiller Number (**TN**)	Mean Metaxylem Size (**MMXS**-mm^2)
14	Nodal Root Number (**RN**)	Metaxylem Number (**MXN**)
15	Shoot Dry weight (**SD**-g)	Lacunae Number (**LAN**)
16	Root Dry weight (**RD**-g)	–
17	Root Shoot Ratio (**RSRatio**)	–
18	Stomatal Conductance (**CD**)	–
19	Chlorophyll Content Index (**CCI**)	–
20	Photo Synthesis (**PTS**)	–

These traits were subsequently used in a machine-learning model to identify which traits might be most responsible for those classifications.

emission that have been reported from rice fields, (i) at maximum tillering and (ii) at panicle initiation to flowering (Gogoi et al. 2005, Islam et al. 2019, Wang et al. 2021, Tamilmounika et al. 2024). Since our experiment was directly seeded after germination, we expected the plants to reach maximum tillering stage around 45 days after sowing. Therefore the sampling of root and shoot tissue was conducted at 45 days after sowing. The shoots of all plants were sampled by cutting them from the base of the plant and placing them in paper bags. After collecting the shoots, the pots containing roots and soil were placed under gently running tap water to wash away the soil. The roots were handled carefully during washing to avoid any damage. The six longest intact nodal roots extending from base to apex were sampled from each plant. These were divided into two plastic bags, with three roots in each, and preserved in 50% ethanol. The remaining roots were also preserved in 25% ethanol in a plastic bag. All root samples were stored in a cold room (4°C) until further processing.

Nodal root number was counted in the laboratory after washing the roots. The earlier preserved three roots from each plant were scanned to obtain an 8-bit greyscale image using an Epson Perfection 7000 scanner at a resolution of 600 dots per inch. All lateral roots were well spread out in the scanning tray so that the number of tips could be detected. The images were analysed using the root system analysis software (WinRHIZO 2004b, Regent Instruments, Inc., Quebec, Canada) to generate other root morphological traits (Table 2).

The remaining three nodal roots were used to measure the root porosity at different positions: base root porosity (BP), middle root porosity (MP), and tip root porosity (TP). We used 3-cm segments (with lateral roots excised) from the nodal root base, middle, and tip to measure the root porosity. Three-centimetre segments were taken 1 cm away from the base for BP, from the midpoint of the root for MP, and 0.5 cm away from the tip for TP. Each segment was cut into approximately four equal parts to fit into a 0.5 ml Eppendorf microfuge tube for vacuum infiltration. Porosity was measured using the microbalance method (Visser and Bögemann 2003). The six roots selected for morphological and anatomical analysis were then combined with the remaining root samples. Shoots and roots from each plant were then dried in an oven at 70°C for 5 days to determine shoot dry weight (SD, g) and root dry weight (RD, g).

2.4 Anatomical traits

Fifteen anatomical traits (Table 2) were taken from two roots per plant from two replicate plants. Nodal roots (two roots per plant) were cross-sectioned by hand at 5 cm from the base into thin sections using a surgical blade under a stereomicroscope. The sections were stained with 1% (v/v) acridine orange for 20–25 s and cleaned with 2–3 drops of clearing solution (70% autoclaved glycerol). Each cross-section was imaged under a compound microscope with a digital camera (Nikon DS-Ri2, Nikon, Nikon Instruments Inc. Melville, NY, USA) under 5× and 10× objectives. The cross-section images were analysed by the anatomical trait analysis software RootScan.2.4 (Burton et al. 2012). Ten images per replication, for a total of 500 images, were analysed from 50 plants. Anatomical trait computation details are provided in Supplementary Table S2, Figs S2 and S3. ANOVA and correlation analysis of root porosity, morphological, and anatomical

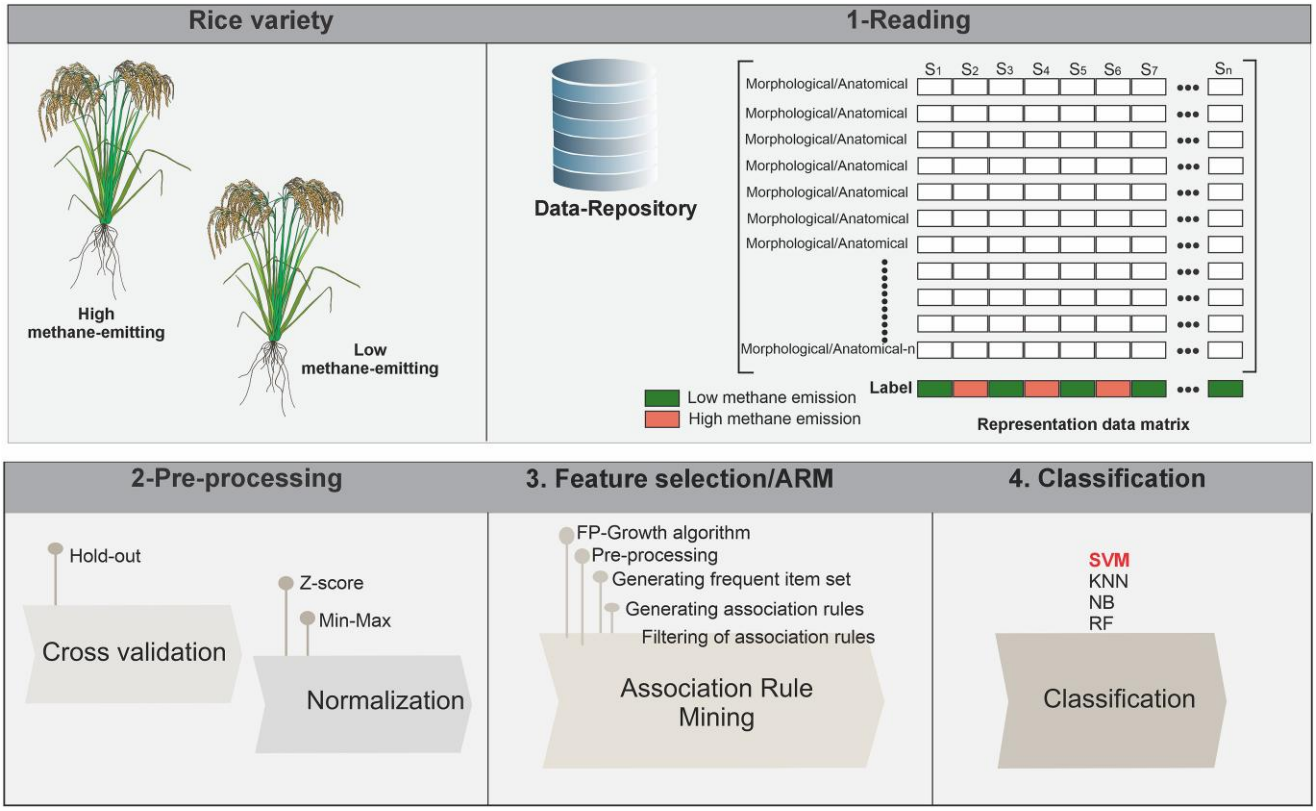

Figure 1. Overview of the development of a machine-learning model to pinpoint the key contrasting traits between high and low methane-emitting rice varieties. Four main steps, composed of reading, pre-processing, feature selection/association rule mining, and classification were applied to morphological and anatomical features. (1) In the reading step, morphological and anatomical results from a greenhouse study were used. (2) The pre-processing step encompassed two sub-steps, cross-validation and normalization. (3) The feature selection step was employed by association rule mining approach on morphological and anatomical features, separately, and sort the features based on their impact on the high/low methane emission labels. (4) The SVM classifier model was utilized to classify high and low methane emission features based on the provided features from the previous step.

traits were conducted in R v. 4.4.2 (R Core Team 2024) using the libraries *agricolae* and *corrplot*.

2.5 Machine learning procedures

The development of the machine-learning model involved four key steps: loading/reading, pre-processing, feature selection, and classification, as illustrated in Fig. 1. Within the first and second steps, morphological and anatomical datasets were individually pre-processed. During the feature selection step, irrelevant and redundant traits were filtered out using the association rule mining algorithm (Agrawal et al. 1993) allowing for the identification of the most significant features that differentiate high and low methane-emitting rice varieties. These selected features were then used to train a classification model to distinguish between the two groups. Python v. 3.11.9 served as the primary programming language, using libraries Numpy, Matplotlib, Pandas, MLXtend, and Scikit-learn to build and implement the machine learning algorithms, as well as to generate the related figures.

2.6 Loading/reading and pre-processing

Separate pre-processing procedures were applied to the morphological and anatomical datasets, which included data organization, cross-validation, and normalization. Initially, the datasets were structured into matrices with rows corresponding to features and columns corresponding to samples. Next, the hold-out cross-validation approach was utilized to divide the data into training and testing subsets. For normalization, both z-score standardization and min–max scaling were implemented using Equations 1 and 2, respectively. These normalization techniques were selectively applied during the feature selection and classification steps, depending on the specific requirements of the algorithms employed.

$$y = \frac{1}{\text{Max} - \text{Min}}(x - \text{Min}) \quad (1)$$

$$y = \frac{x - \mu}{\sigma}. \quad (2)$$

In Equations 1 and 2, x represents the original feature value, while y denotes the normalized value. The parameters μ and σ correspond to the mean and standard deviation of x, respectively, whereas Min and Max indicate the minimum and maximum values of x, in that order.

2.7 Feature selection

Feature selection (FS) is a crucial step in machine learning and pattern recognition, as it can enhance classification accuracy or

reduce regressor error (Ferreira and Figueiredo 2012). Since some features are irrelevant and redundant in real data concerning a given learning task, it is useful to filter out these features by extracting the most relevant subset of features. This process can help avoid over-fitting and accelerate the training process (Bommert et al. 2020). In this study, we applied association rule mining as the feature selection technique, treating morphological and anatomical data independently. This method enabled us to isolate strongly discriminative features, uncovering potential patterns and relationships among variables, which the classifier model later used for effective group differentiation.

2.8 Association rule mining

Association rule mining is a powerful technique used to uncover meaningful relationships among items within a dataset. An association rule is generally expressed in the form $A \rightarrow C$, where A (Antecedent) and C (Consequent) represent conditions or outcomes. When A is treated as one or more features and C as either additional features or a user-defined phenotype, the method can reveal significant associations between features or between features and phenotypes. This makes it a suitable approach for identifying key features from a broader set of candidates. In our study, the association rule mining algorithm was applied to the candidate feature set. To support this process, we utilized fundamental metrics such as Support, Confidence, and Lift which are central to evaluating the strength and relevance of the discovered rules. Let $I = \{i_1. i_2. \ldots i_d .y\}$ be a set of items, $D = \{d_1. d_2. \ldots . d_n\}$ be a dataset of n instances, $F = \{f_1. f_2. \ldots . f_m\}$ be the features space with m features, and $Y = \{0. 1\}$ be the user-defined phenotype. The d_i can be presented as a tuple $(X_i. y_i)$, where $X_i \in f_1 \times f_2 \times \ldots \times f_m$ and $y_i \in Y$. Also, $A \rightarrow C$ is an association rule, where $A \subset I$, $C \subset I$, and $A \cap B = \varphi$. The support of rule $A \rightarrow C$ is a probability of instances contain both A and C, as shown in Eq. 3. Support evaluates the rule's usefulness.

$$\text{Support}(A \rightarrow C) = \frac{\text{Support}(A \cup C)}{n}. \quad (3)$$

In Equation 4, $\text{Support}(A) = |\{ d_i | A \subseteq X_i. d_i \in D\}|$ is the number of instances that includes the i itemset. The confidence of rule $A \rightarrow C$ is the probability, which shows the frequency of cases with C among all samples that contain A, as shown in Eq. 4. Confidence estimates the rule's certainty.

$$\text{Confidence}(A \rightarrow C) = P(C|A) = \frac{\text{Support}(A \cup C)}{\text{Support}(A)}. \quad (4)$$

The Lift of rule $A \rightarrow C$ determines the dependency between the occurrence of itemset A and C. When the Lift value is more (less) than one, the occurrence of A is positively (negatively) correlated with the occurrence of C. If the Lift value is equal to one, then A and C are independent. The Lift value is shown in Equation 5.

$$\text{Lift}(A \rightarrow C) = \frac{P(A \cup C)}{P(A)P(C)}. \quad (5)$$

Association rule mining typically involves two primary steps: generating frequent itemsets and creating association rules. In the first step, the algorithm iteratively identifies itemsets and those with support values exceeding a predefined threshold (min_support) are considered frequent. The second stage focuses on generating association rules based on these frequent itemsets. One of the earliest algorithms to effectively implement this approach was the Apriori algorithm, introduced by Agrawal et al. (1996), which utilizes the principle of frequent itemsets for rule extraction.

In the present study, we adopted the FP-Growth algorithm for association analysis, which is widely regarded as an efficient approach due to its ability to traverse the itemset space in a single pass, offering advantages in computation time and memory efficiency (Han et al. 2004). Detailed pseudo-code for the FP-Growth algorithm is presented in Tables 3 and 4.

2.9 Classification step

Classification is a fundamental component of many machine learning applications, where the model is trained to differentiate between data categories based on selected input features. Distinction is achieved by establishing either linear or nonlinear boundaries. The effectiveness of the classification model reflects how well it separates the defined classes. Strong classifier model performance is indicated by the chosen features possessing meaningful discriminative power—in the case of this study, for identifying methane emission levels—the classification accuracy is not negatively impacted by removing less relevant features (Lecun et al. 2015). For this purpose, we utilized a robust classification algorithm—SVM—due to its effectiveness in handling highly dimensional data and complex boundaries.

Initially, the SVM model was trained using the designated training dataset. Following the training step, the model's classification performance was estimated: classification accuracy and the area under the receiver operating characteristic curve (AUC-ROC) were defined in Equation 6.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}} \times 100 \quad (6)$$

where TP, TN, FP, and FN are true positive, true negative, false positive, and false negative, respectively. AUC-ROC is the area under the receiver operating characteristic (ROC) curve.

3. RESULTS

Throughout the greenhouse experiment, the mean temperature was recorded at 26.8°C, while the average humidity was 76.9%, and all varieties grew well under those conditions (Supplementary Fig. S1). Among the locations along the root axis from which root porosity (RP) was determined, a greater degree of porosity was seen at the nodal root base and middle section and least at the root tip (Fig. 2A). However, all porosity-related traits (porosity, root diameter, % aerenchyma, and number of lacunae) were positively correlated (Fig. 2B).

The machine learning workflow consisted of four core stages: data loading, pre-processing, feature selection, and classification (Fig. 1). During processing, 20 morphological traits from 75 samples were organized into a matrix with 20 rows and 75 columns. Anatomical data comprised 15 traits obtained from 50 samples and were arranged in matrix format as well. Both datasets were split into training and testing subsets using the hold-out cross-validation method, with a 70%:30% ratio. To

Table 3. Pseudocode of frequent itemset generation step in FP-Growth algorithm.

Algorithm: Frequent itemset generation in FP-Growth algorithm
Input: A database DB, represented by FP-tree constructed, and a minimum support threshold ξ.
output: The complete set of frequent patterns
Method: FP-growth (Tree, α)
(1) **if** Tree contains a single prefix path, **then**://Mining single prefix-path FP-tree
(2) let P be the single prefix-path part of Tree;
(3) let Q be the multipath part with the top branching node replaced by a null root;
(4) **for** each combination (denoted as β) of the nodes in the path P **do**:
(5) generate pattern $\beta \cup \alpha$ with support = minimum support of nodes in β;
(6) let freq_pattern_set(P) be the set of patterns so generated;
(7) **end for**
(8) **else** let Q be Tree:
(9) **for** each item ai in Q **do**://Mining multipath FP-tree
(10) generate pattern $\beta = a_i \cup \alpha$ with support = a_i.support;
(11) construct β's conditional pattern-base and then β's conditional FP-tree Treeβ;
(12) **if** Treeβ = ∅ **then**:
(13) call FP-growth (Treeβ, β);
(14) let freq_pattern_set(Q) be the set of patterns so generated;
(15) **end if**
(16) **end for**
(17) **end if**
(18) return (freq_pattern_set(P) ∪ freq_pattern_set(Q) ∪ (freq_pattern_set(P) × freq_pattern_set(Q)))

Table 4. Pseudocode of rules generation step in FP-Growth algorithm.

Algorithm: Rules generation in FP-Growth algorithm
Method:
(1) **For** each frequent itemset k-itemset F_k, $k \geq 2$ **do:**
(2) $H_1 = \{i | i \in F_k\}$;//1-item consequents of the rule.
(3) call rules_generator (F_k, H_1);
(4) **End for**
Procedure rules_generator (F_k, H_m)
(1) $k = |F_k|$;//size of frequent itemset
(2) $m = |H_m|$;//size of rule consequent
(3) **if** $k > m + 1$ **then:**
(4) $H_m = m + 1 -$ item consequents generated from H_m;
(5) **For** each $h_{m+1} \in H_{m+1}$ **do:**
(6) Confidence = $\frac{\text{Support}(F_k)}{\text{Support}(F_k - H_{m+1})}$;
(7) **if** Confidence ≥ min_confidence **then:**
(8) output: the rule $(F_k - h_{m+1}) \rightarrow h_{m+1}$;
(9) **Else:**
(10) delete h_{m+1} from H_{m+1};
(11) **End if**
(12) **End for**
(13) call ap-genrules(F_k, H_{m+1})
(14) **End if**

standardize the data, min–max normalization was applied based on statistics derived from the training set.

Once the pre-processing was complete, feature selection was conducted separately for the morphological and anatomical datasets. Through association rule mining, meaningful patterns were identified, including relationships between features (morphological/anatomical) and between features and the target variable (high/low methane-emitting levels). Significant traits were determined based on their frequency in the generated rules (Figs 6A and 7A), helping to highlight their importance in distinguishing between high and low methane-emitting rice varieties. For the morphological and anatomical datasets, five key features associated with classification as high or low methane-emitting were identified. The association rules determining that their consequent part is equal to classification as high/low methane-emitting were called high/low methane-emitting association rules. All morphological/anatomical features resulting from low and high methane-emitting rules based on repeat counts are shown in Supplementary Table S3. The algorithm parameters were set to a minimum-support threshold of 0.1 and a lift value of 1.1.

Spearman correlation analysis identified distinct morphological and anatomical traits that significantly differentiated high and low methane-emitting rice varieties (Figs 3A and

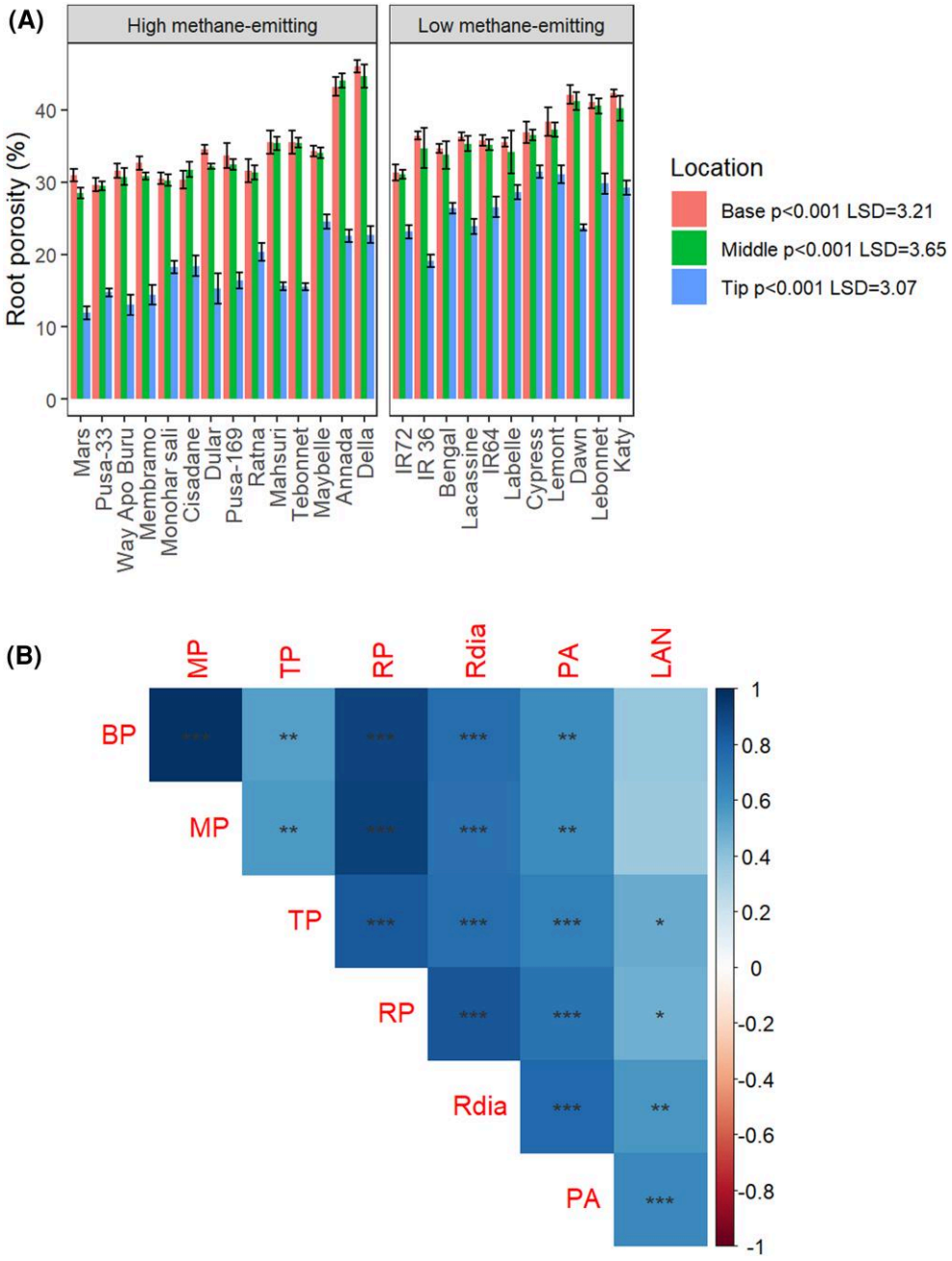

Figure 2. Root porosity-related phenotypic data from the greenhouse experiment. (A) the nodal root porosity at the base, middle and tip. Mean values (±s.e.) and the *P*-value and LSD for the varietal effect at each location are shown. (B) Correlations among porosity-related traits. Trait abbreviations are given in Table 2. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

4A). The robustness of these correlations was further supported through rule strength metrics-support, lift, and confidence-derived from both morphological and anatomical datasets (Figs 3B and 4B). High methane-emitting varieties were strongly associated with specific morphological traits, including lower MP, BP, RP, and Rdia values, and higher STLR. Similarly, high methane-emitting varieties were strongly linked to anatomical traits such as reduced LAN, TSA, MMXS, MXN, and MXVA. In contrast, low methane-emitting varieties were closely associated with morphological traits such as lower RN, TN, RD, and LN, along with higher RDia. Low methane-emitting varieties also showed strong associations with anatomical traits, including elevated values of PA, TSA, RCXS, RSXS, and AA. These associations were prominently visualized in network graphs, where trait connections reflected their frequency in rules associated with methane emission levels (Figs 3C/D and 4C/D).

During the classification step, the selected morphological and anatomical traits were evaluated for their ability to distinguish between high and low methane-emitting groups. The strong associations between morphological and anatomical traits and methane emission levels were further supported by box plot analyses, where comparisons of medians, interquartile ranges, and whiskers revealed clear distinctions between high

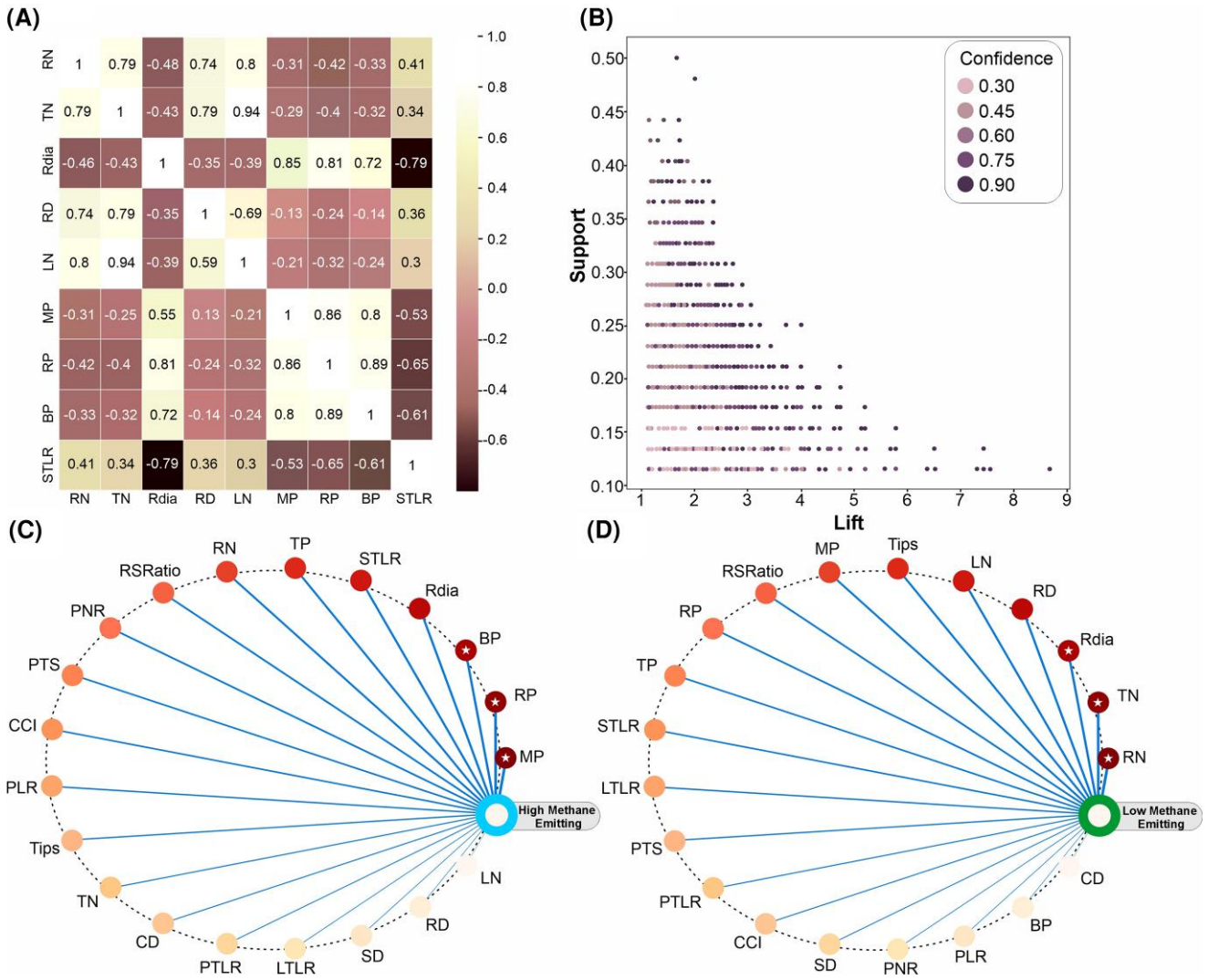

Figure 3. (A) Heatmap plot based of Spearman correlations among significant morphological features (B) association rules according to their support, lift, and confidence. Graph network of top morphological features with high repeat count in (C) high methane-emitting and (D) low methane-emitting association rules. The most frequent morphological features were displayed with the weight edge based on repeat counts. Darker colours of the circle (C and D) indicate a stronger correlation with classification as either high or low methane-emitting.

and low methane-emitting rice groups (Fig. 5). These findings were reinforced by violin plots, which consistently highlighted the same five key morphological and anatomical traits differentiating high and low methane-emitting varietal groups (Supplementary Fig. S4).

A SVM model with both radial basis function (rbf) and linear kernels was employed to assess the discriminative power of these features. The selected morphological features could classify the high and low methane-emitting varietal groups in the train/test set with high accuracy (98%/100%) and AUC (0.98/0.95). For the anatomical traits, the selected features could also classify the high and low methane-emitting varietal groups in the train/test sets with high accuracy (89%/93%) and AUC (0.94/0.90). The confusion matrix and ROC curve of the train/test sets are illustrated in Figs 6B,C and 7B,C, respectively. The pair plot of top morphological and anatomical features is illustrated in Supplementary Figs S5 and S6, respectively.

Among morphological features, MP and RN were the most frequent itemset with 2246 and 4962 repeat counts in the high and low methane-emitting association rules, respectively. As a result, we used other association rules to investigate their relations with other features. Based on association rules in the graph network, it is obvious that lower MP, the most frequent feature in the high methane association rules, had a high dependency on BP, RP, RN, PTS, STLR, Rdia, TP, and RSRatio (Fig. 8A). RN, the most frequent feature in the low methane association rules, had a high dependency on RDia, TN, MP, RP, STLR, TP, LN, RSRatio, Tips, and RD (Fig. 8B).

Among anatomical features, lower values of LAN and higher values of PA were the most frequent itemset with 1420 and 90 repeat counts in high and low methane-emitting association rules, respectively. Lower values of LAN, the most frequent feature in the high methane-emitting association rules, had a high dependency on PLCC, TSA, AA, RCS, MMXS, MXN, and MXVA (Fig. 9A). Higher values of PA, the most frequent feature in the low methane-emitting association rules, had a high dependency on TCA, RXSA, AA, RSXS, RCXS, TSA, PMXVAS, MXN, PLCC, and LAN (Fig. 9B).

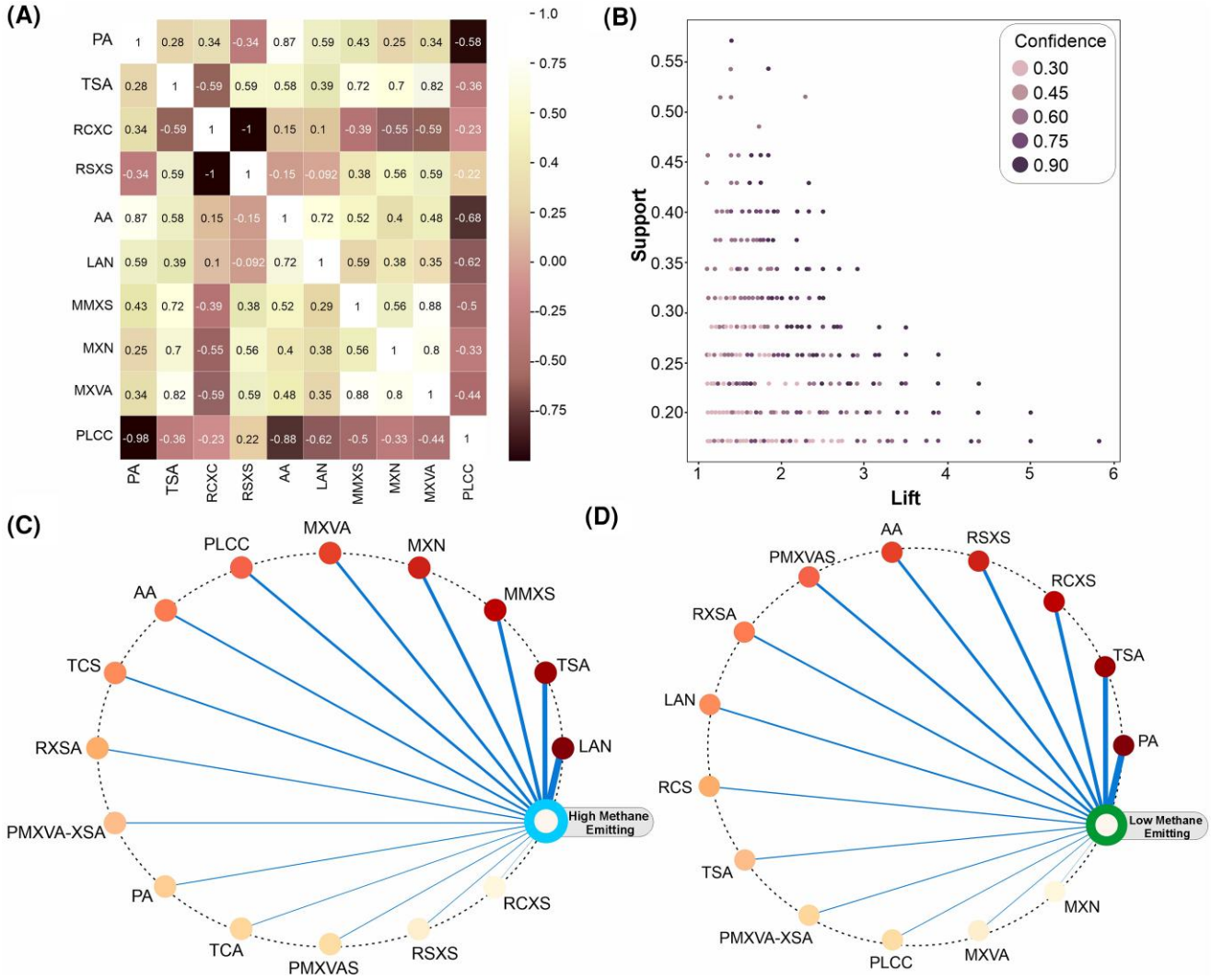

Figure 4. (A) Heatmap plot based of the Spearman correlations among significant anatomical features (B) association rules according to their support, lift, and confidence. Graph network of top anatomical features with high repeat count in (C) high methane-emitting and (D) low methane-emitting association rules. The most frequent anatomical features were displayed with the weight edge based on repeat counts. Darker colours of the circle (C and D) indicate a stronger correlation with classification as either high or low methane-emitting.

4. DISCUSSION

This study combined morphological and root anatomical characterization of rice varieties that had been previously classified as high or low methane-emitting, with machine learning approaches to identify which traits might be most responsible for those classifications. Through Association Rule Mining, 10 morphological and 10 anatomical traits emerged as the most frequently recurring features, suggesting their strong relevance in differentiating between high and low methane-emitting varieties. A selected set of morphological traits successfully distinguished the high from low methane-emitting varieties in train/test data with high accuracy in both training and testing datasets (95% and 98%, respectively; Fig. 6C). Likewise, a set of anatomical features achieved strong classification performance, with accuracy of 94% for training data and 90% for testing data (Fig. 7C).

The traits and correlations identified by association rule mining analysis allowed us to characterize the high and low methane-emitting varieties into general plant types. High methane-emitting varieties were characterized as having less porous, smaller diameter nodal roots, and more fine lateral roots. Low methane-emitting varieties were characterized as having fewer nodal roots and tillers, and larger-diameter nodal roots with greater aerenchyma formation.

Our analysis is based on several methodological assumptions that warrant consideration when interpreting the results. A central assumption is that the varietal classifications of low and high methane-emitting from the previous studies are robust. These varietal classifications have generally not undergone extensive confirmation, and the varieties classified as low and high methane-emitting were not necessarily identified from the same study (Table 1, Supplementary Table S1). Therefore it may have been preferable to validate the varietal classifications of previous studies through multi-season field trials prior to selection. However, there were several advantages of our approach: more varieties could be included than in a typical methane emissions field trial, we were able to take advantage of existing

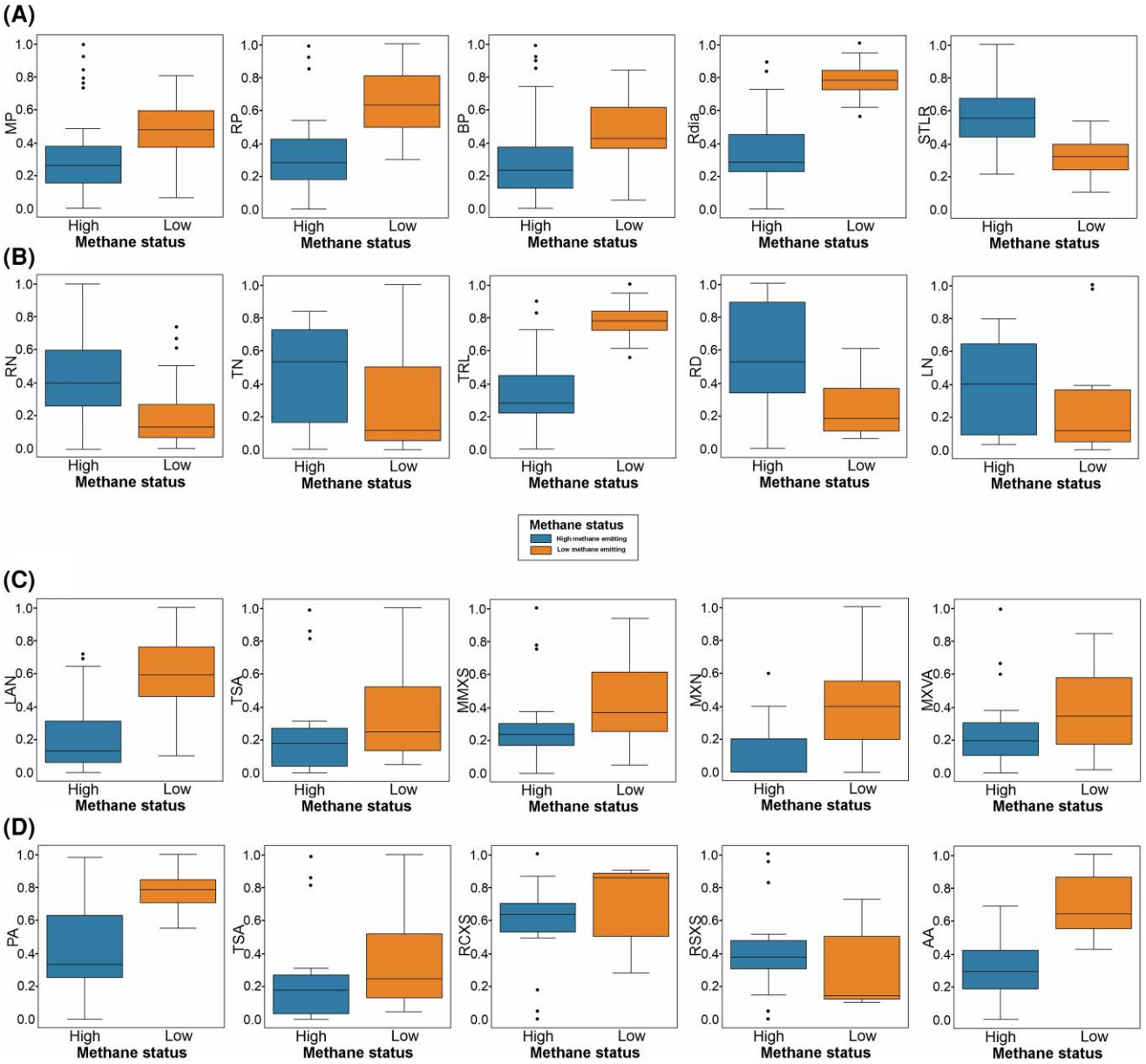

Figure 5. Boxplots of selected morphological features for (A) high and (B) low methane-emitting variety classes. Boxplots of selected anatomical features for (C) high and (D) low methane-emitting variety classes. Features values of morphological and anatomical were from 0 to 1 by the min–max method.

data, and varieties classified as high- or low-emitting across different environments could be included.

We also assume that the root measurements collected during the greenhouse experiment are representative of the actual root systems present in the field at the time of methane emissions monitoring. This assumption is particularly important, as practical constraints often prevent the simultaneous measurement of both root traits and gas fluxes. These assumptions support the application of a machine learning approach in our study. By leveraging complex, multivariate patterns in the data, machine learning models enable us to explore nonlinear relationships between root traits and methane emissions that may be obscured by traditional statistical techniques. However, the strength of our model's predictions is inherently tied to the validity of the aforementioned assumptions. Furthermore, there may be interesting interactions between anatomical and morphological features that affect methane emissions, for example between nodal root length (depth of penetration into the soil) and porosity (e.g. Jiménez and Pedersen 2023). Future work could incorporate a combined model to evaluate such possible effects.

4.1 Morphological traits

All five identified morphological features for high methane-emitting varieties were directly or indirectly related to average root diameter (RDia) and root porosity (RP). RDia is the average diameter of all L-type lateral roots (LTLR), S-type lateral roots (STLR), and nodal roots (RN). Although their diameter is small, fine lateral roots have a large effect on average root

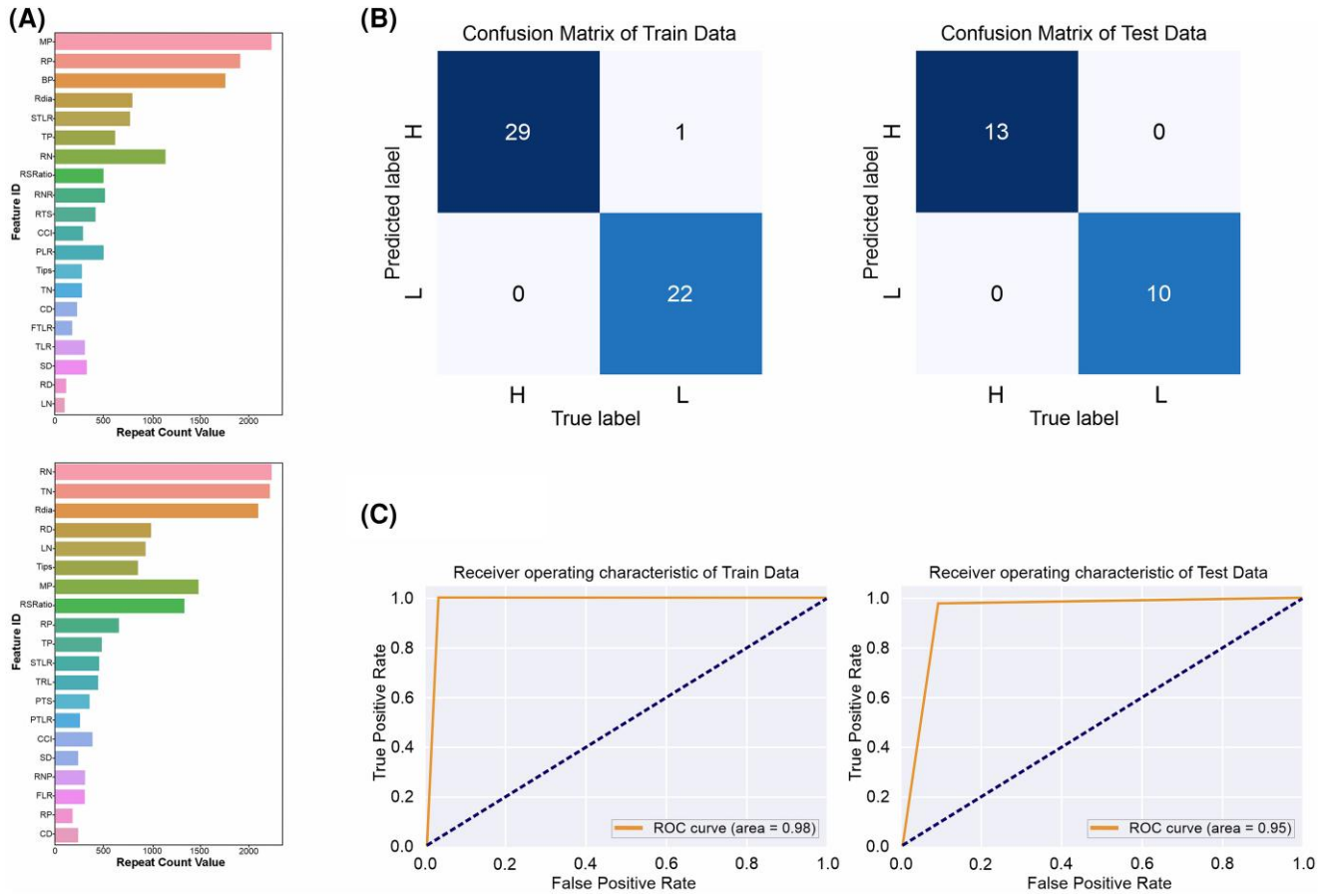

Figure 6. The performance of feature selection and classification steps in the morphological data. (A) Bar plot of candidate morphological high/low methane-emitting features based on repeat count values. (B) Confusion matrix of training and test data. (C) ROC curve of training and test data. The area under the curve of receiver operating characteristic (AUC-ROC) is reported for each curve.

diameter since the length of lateral roots is typically much greater than that of the nodal roots (Yanai et al. 2013). Therefore, the percent S-type lateral roots (STLR) was negatively correlated with average Rdia. RP was positively correlated with RDia, BP, and MP as well as BP and MP positively associated with each other (Fig. 3A). STLR was also negatively correlated with BP, MP, and RP (Fig. 3A), which is in contrast with previous reports of aerenchyma promoting lateral root formation (Niones et al. 2012). However, although the porosity measurements in this study did not include lateral roots, a lower degree of aerenchyma formation has been observed in S-type lateral roots compared to L-type lateral roots and nodal roots (Noorrohmah et al. 2020).

Root porosity (RP) is the ratio of the air-filled volume within the root to the total volume, and is primarily influenced by parameters such as root diameter and the extent of aerenchyma development (Justin and Armstrong 1987, Colmer 2003a). Variation in porosity along the nodal root length reflects the pattern of aerenchyma formation, which typically begins at the apical regions and progressively extends tower to the base (Ranathunge et al. 2003), reaching full development at the basal part (Armstrong and Armstrong 1994, Kozela and Regan 2003).

Our model predicted low RP (MP) to be most strongly associated with high methane-emitting varieties (Fig. 8A). A more porous root might reduce methane emissions from the rice field by increasing the capacity to hold and carry gases, boosting O_2 pressure to the root tip to diffuse into the rhizosphere. which is constrained by the well-structured barriers to radial O_2 loss (Ejiri and Shiono 2019). Thus, root porosity has a robust role in radial oxygen loss into the rhizosphere (Luxmoore et al. 1970, Gutierrez et al. 2014). O_2 can mediate methane emissions from paddy soil in terms of production and reduction. It increases the availability of electron acceptors (NO_3^-, Fe_3^+, Mn_4^+, SO_4^-) by re-oxidizing the reduced form of methane in the rhizosphere (Armstrong 1970, Trolldenier 1988, Revsbech et al. 1999, Van Der Welle et al. 2006, Yang et al. 2012) due to its higher oxidation potential (Husson 2013). These electron acceptors raise the soil redox potential (Eh), which is toxic to methanogens whose favourable habitat is more reduced (−105 mV) Eh (Wang et al. 1993, Xu et al. 2003). Consequently, methane production becomes hindered by the presence of O_2 in the rhizosphere. In addition, O_2 can create a favourable environment for methane-consuming methanotrophs, which can consume 80%–90% of produced methane at the rhizosphere by oxidizing methane to methanol, formaldehyde, formate, and finally CO_2 (Hanson and Hanson 1996, Bowman 2006), thereby decreasing methane emissions from rice fields. Among the traits measured in this study, a lower MP was ranked as the top trait related to high methane-emitting varieties (Fig. 8A), which may be due to its limited contribution

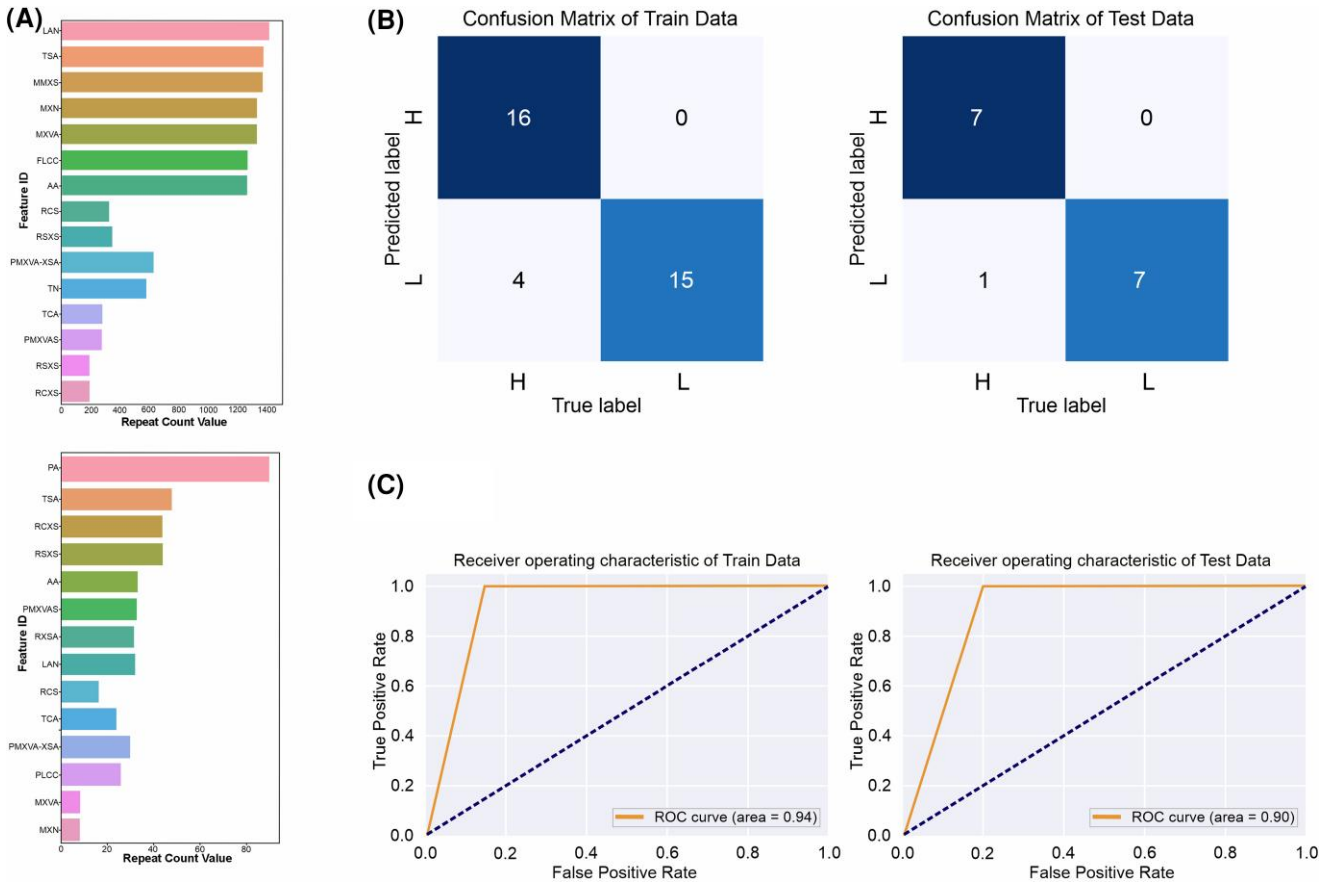

Figure 7. The performance of feature selection and classification steps in the anatomical data. (A) Bar plot of candidate morphological high/low methane-emitting features based on repeat count values. (B) Confusion matrix of training and test data. (C) ROC curve of training and test data. The area under the curve of receiver operating characteristic (AUC-ROC) is reported for each curve.

to root porosity, O_2 holding, and transporting capacity to rhizosphere that demands further validation.

The possible relationship of RD and RN with methane-emitting classification observed here has been reported previously (Aulakh et al. 2001a). RN, which our model identified as most strongly related to low methane-emitting varieties, was positively correlated with TN, RD, and LN (Fig. 3A). It has been reported that aboveground and belowground biomass had a significant role in methane emissions, which may be due to their role in providing substrates for methanogen activity via photosynthesis and root exudation of assimilates. Root exudates are a major carbon source of the rhizobiome, including methanogenic bacteria which produce methane (Qi et al. 2024). The top ranking of lower nodal root number (RN) among traits associated with low methane-emitting varieties may suggest that reduced numbers of roots may contribute to lower methane emissions by limiting the release of root exudates and decreasing the efficiency of methane transported from the soil to the atmosphere (Aulakh et al. 2001b).

4.2 Anatomical traits

The top five anatomical traits identified to be associated with high methane-emitting varieties (LAN, TSA, MMXS, MXN, and MXVA) and the top five traits associated with low methane-emitting varieties (PA, TSA, RCXS, RSXS, and AA) were directly or indirectly related to the identified key morphological traits root porosity (RP) and average root diameter (Rdia) (Supplementary Fig. S2). In addition, these 10 anatomical traits were the secondary variable (Supplementary Table S2) generated from cross-section of the root, which were linked to root diameter (RDia).

Aerenchyma is the main conduit for holding and carrying gases as it is longitudinally developed throughout the root (Colmer 2003b, Wang et al. 2017). It has been noted that the increment of root porosity solely depended on increased root diameter (Striker et al. 2007) as a greater level of aerenchyma area has been observed to develop in larger diameter roots (Kirk 2003, Kim et al. 2018). In this study, aerenchyma area (AA) was significantly and positively correlated with root diameter and root porosity (Fig. 2B), further reflecting the relationship between root diameter (size) and porosity (as influenced by aerenchyma formation). Likewise, stele and metaxylem size often reflect larger root diameters (e.g. Uga et al. 2008), and were positively correlated with PA, AA, LAN (Fig. 4A) in this study.

A well-developed aerenchyma network is known to facilitate the bidirectional transport of gases enabling methane movement from soil to the atmosphere and oxygen diffusion from the atmosphere into the rhizosphere (Wassmann et al. 2000,

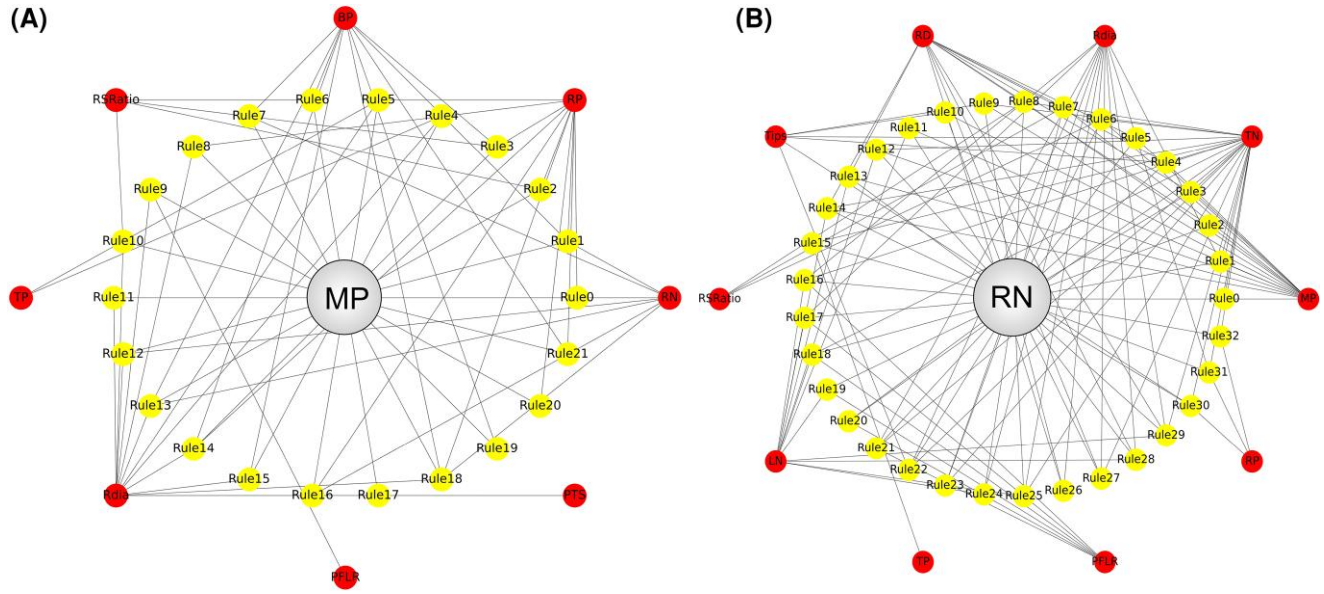

Figure 8. Graph network of high and low methane-emitting rules for morphological features, respectively. (A) MP related association rules (with lift > 1.7 and support > 0.24), in which the identified morphological features, its rules, highest repeat count feature (MP) are presented. (B) RN related association rules (with lift > 2 and support > 0.24), in which the identified morphological feature, its rules, highest repeat count feature (RN) are presented.

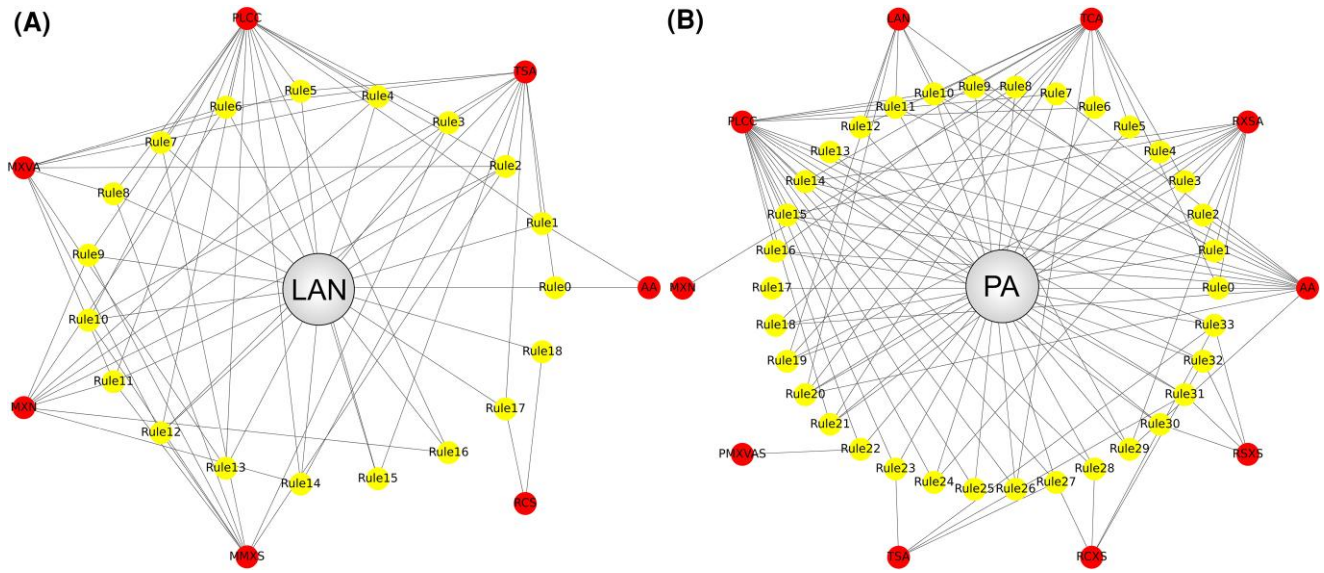

Figure 9. Graph network of high and low methane-emitting rules for anatomical classes, respectively. (A) LAN-related association rules (with lift > 1.7 and support > 0.22), in which the identified anatomical feature, its rules, highest repeat count feature (LAN) are presented. (B) PA-related association rules (with lift > 1.7 and support > 0.29), in which the identified morphological feature, its rules, highest repeat count feature (PA) are presented.

Mei et al. 2009, Win et al. 2012, Wang et al. 2017). In this study, the lower LAN related to high methane-emitting varieties and the higher PA related to low methane-emitting varieties suggests that the role of aerenchyma in rhizosphere oxygenation (decreasing the activity of methanogens) might predominate over the role of aerenchyma in transporting methane from the soil to the atmosphere.

5. CONCLUSIONS

The machine learning method presented here is a novel approach to identify important morphological and anatomical characteristics to differentiate high/low methane-emitting varieties. We identified trait patterns that appear to be related to varieties previously reported to be high or low methane-emitting. Our methodology separately identified 10 top morphological and anatomical

characteristics with the highest effect on the varietal classification of high or low methane-emitting. High methane-emitting varieties were characterized as having less porous, smaller diameter nodal roots and more fine lateral roots. Low methane-emitting varieties were characterized as having fewer nodal roots and tillers, and larger-diameter nodal roots with greater aerenchyma formation. These findings indicate that root traits such as porosity, aerenchyma development and higher diameter may play a more critical role in enhancing rhizosphere oxygenation than in facilitating methane transport from soil to the atmosphere. This study serves as an initial step in applying machine learning approach to explore the underlying mechanisms driving methane emission in rice cultivation systems and varieties, since several assumptions were made regarding the robustness of the variety classification and the use of greenhouse-grown roots to represent field-grown roots. The models implemented here can help deal with the inconsistencies and large degree of variability related to methane and root measurements in rice paddy fields.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the valuable suggestions and kind assistance provided by the members of the Plant Molecular Biology Laboratory and the Plant Physiology Laboratory at the International Rice Research Institute (IRRI), Philippines. We thank Dr. Tao Li for constructive comments on the manuscript.

AUTHOR CONTRIBUTIONS

R.K.R.: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing–original draft, Writing–review & editing. S.M.H.K: Conceptualization, Data curation, Formal Analysis, Software, Visualization, Writing—review & editing. A.K.: Conceptualization, Funding acquisition, Project administration, Supervision, Resource, Project administration, Writing–review & editing. A.H.: Conceptualization, Funding acquisition, Supervision, Validation, Resource, Project administration, Writing–review & editing. S.R.K.: Conceptualization, Supervision, Validation. R.K.T.: Conceptualization, Supervision, Validation. S.P.: Formal Analysis, Validation. J.E.H.: Conceptualization, Supervision, Validation, Writing–review & editing. M.G.Q.D.: Conceptualization, Supervision, Validation, Writing–review & editing. E.T.M.O.: Conceptualization, Supervision, Validation, Writing—review & editing.

SUPPLEMENTARY DATA

Supplementary data is available at *in silico Plants* online.

Conflict of interest: The authors declare that they have no competing interests.

FUNDING

This study was partially funded by the ICAR-IRRI Program and National Agricultural Technology Project–Phase II (NATP-2), Farmgate, Dhaka, Bangladesh.

DATA AVAILABILITY

The data supporting the findings of this study are included in the main article and its Supplementary Information files. The varieties used in this study are available from the IRRI Genebank, Philippines.

GENERATIVE AI STATEMENT

The author(s) affirm that no generative AI tools were utilized in the writing or preparation of this manuscript.

REFERENCES

Adhya TK, Rath AK, Gupta PK *et al.* Methane emission from flooded rice fields under irrigated conditions. *Biol Fertil Soils* 1994;**18**:245–8. https://doi.org/10.1007/BF00647675

Agrawal R, Imieliński T, Swami A. Mining association rules between sets of items in large databases. *ACM SIGMOD Rec* 1993;**22**:207–16. https://doi.org/10.1145/170036.170072

Agrawal R, Mannila H, Srikant R *et al.* Fast discovery of association rules. *Adv Knowl Discov Data Min* 1996;**12**:307–328.

Armstrong W. Rhizosphere oxidation in rice and other species: a mathematical model based on the oxygen flux component. *Physiol Plant* 1970;**23**:623–30. https://doi.org/10.1111/j.1399-3054.1970.tb06455.x

Armstrong J, Armstrong W. Chlorophyll development in mature lysigenous and schizogenous root aerenchymas provides evidence of continuing cortical cell viability. *New Phytol* 1994;**126**:493–7. https://doi.org/10.1111/j.1469-8137.1994.tb04246.x

Aulakh MS, Bodenbender J, Wassmann R *et al.* Methane transport capacity of rice plants. II. Variations among different rice cultivars and relationship with morphological characteristics. *Nutr Cycl Agroecosyst* 2000;**58**:367–75. https://doi.org/10.1007/978-94-010-0898-3_32

Aulakh MS, Wassmann R, Bueno C *et al.* Characterization of root exudates at different growth stages of ten rice (*Oryza sativa* L.) cultivars. *Plant Biol* 2001a;**3**:139–48. https://doi.org/10.1055/s-2001-12905

Aulakh MS, Wassmann R, Rennenberg H. Methane emissions from rice fields—quantification, role of management, and mitigation options. *Adv Agron* 2001b;**70**:193–260. https://doi.org/10.1016/S0065-2113(01)70006-5

Baruah KK, Gogoi B, Gogoi P. Plant physiological and soil characteristics associated with methane and nitrous oxide emission from rice paddy. *Physiol Mol Biol Plants* 2010;**16**:79–91. https://doi.org/10.1007/s12298-010-0010-1

Bharali A, Baruah KK, Gogoi N. Methane emission from irrigated rice ecosystem: relationship with carbon fixation, partitioning and soil carbon storage. *Paddy Water Environ* 2017a;**15**:221–36. https://doi.org/10.1007/s10333-016-0541-3

Bharali A, Baruah KK, Gogoi N. Potential option for mitigating methane emission from tropical paddy rice through selection of suitable rice varieties. *Crop Pasture Sci* 2017b;**68**:421–33. https://doi.org/10.1071/CP16228

Bommert A, Sun X, Bischl B *et al.* Benchmark for filter methods for feature selection in high-dimensional classification data. *Comput Stat Data Anal* 2020;**143**:106839. https://doi.org/10.1016/j.csda.2019.106839

Bowman J. The methanotrophs—the families Methylococcaceae and Methylocystaceae. In: *The Prokaryotes*. New York, NY: Springer New York, 2006, 266–89.

Burton AL, Williams M, Lynch JP *et al.* RootScan: software for high-throughput analysis of root anatomical traits. *Plant Soil* 2012;**357**:189–203. https://doi.org/10.1007/s11104-012-1138-2

Butterbach-Bahl K, Gasche R, Breuer L *et al.* Fluxes of NO and N2O from temperate forest soils: impact of forest type, N deposition and of liming on the NO and N2O emissions. *Nutr Cycl*

Agroecosystems 1997a;**48**:79–90. https://doi.org/10.1023/A:1009785521107

Butterbach-Bahl K, Papen H, Rennenberg H. Impact of gas transport through rice cultivars on methane emission from rice paddy fields. *Plant Cell Environ* 1997b;**20**:1175–83. https://doi.org/10.1046/j.1365-3040.1997.d01-142.x

Cao M, Gregson K, Marshall S *et al.* Global methane emissions from rice paddies. *Chemosphere* 1996;**33**:879–97. https://doi.org/10.1016/0045-6535(96)00231-7

Chen Y, Li S, Zhang Y *et al.* Rice root morphological and physiological traits interaction with rhizosphere soil and its effect on methane emissions in paddy fields. *Soil Biol Biochem* 2019;**129**:191–200. https://doi.org/10.1016/j.soilbio.2018.11.015

Colmer TD. Aerenchyma and an inducible barrier to radial oxygen loss facilitate root aeration in upland, paddy and deep-water rice (*Oryza sativa* L.). *Ann Bot* 2003a;**91**:301–9. https://doi.org/10.1093/aob/mcf114

Colmer TD. Long-distance transport of gases in plants: a perspective on internal aeration and radial oxygen loss from roots. *Plant Cell Environ* 2003b;**26**:17–36. https://doi.org/10.1046/j.1365-3040.2003.00846.x

Das K, Baruah KK. A comparison of growth and photosynthetic characteristics of two improved rice cultivars on methane emission from rainfed agroecosystem of northeast India. *Agric Ecosyst Environ* 2008a;**124**:105–13. https://doi.org/10.1016/j.agee.2007.09.007

Das K, Baruah KK. Association between contrasting methane emissions of two rice (*Oryza sativa* L.) cultivars from the irrigated agroecosystem of northeast India and their growth and photosynthetic characteristics. *Acta Physiol Plant* 2008b;**30**:569–78. https://doi.org/10.1007/s11738-008-0156-4

Das K, Baruah KK. Methane emission associated with anatomical and morphophysiological characteristics of rice (*Oryza sativa*) plant. *Physiol Plant* 2008c;**134**:303–12. https://doi.org/10.1111/j.1399-3054.2008.01137.x

Denier van der Gon HAC, Neue H-U. Oxidation of methane in the rhizosphere of rice plants. *Biol Fertil Soils* 1996;**22**:359–66. https://doi.org/10.1007/BF00334584

Ding A, Willis CR, Sass RL *et al.* Methane emissions from rice fields: effect of plant height among several rice cultivars. *Global Biogeochem Cycles* 1999;**13**:1045–52. https://doi.org/10.1029/1999GB900084

Duyen DV, Kwon Y, Kabange NR *et al.* Novel QTL associated with aerenchyma-mediated Radial Oxygen Loss (ROL) in rice (*Oryza sativa* L.) under iron (II) sulfide. *Plants* 2022;**11**:788. https://doi.org/10.3390/plants11060788

Edenhofer O, Pichs-Madruga R, Sokona Y *et al.* 2014. *Climate Change 2014: Mitigation. Technical Summary*. https://www.researchgate.net/publication/262048938_Climate_Change_2014_Mitigation_Technical_Summary/references (5 May 2024, date last accessed).

Ejiri M, Shiono K. Prevention of radial oxygen loss is associated with exodermal suberin along adventitious roots of annual wild species of Echinochloa. *Front Plant Sci* 2019;**10**:1–13. https://doi.org/10.3389/fpls.2019.00254

Fang K, Dai W, Chen H *et al.* The effect of integrated rice–frog ecosystem on rice morphological traits and methane emission from paddy fields. *Sci Total Environ* 2021;**783**:147123. https://doi.org/10.1016/j.scitotenv.2021.147123

Ferreira AJ, Figueiredo MAT. Efficient feature selection filters for high-dimensional data. *Pattern Recognit Lett* 2012;**33**:1794–804. https://doi.org/10.1016/j.patrec.2012.05.019

Ghani Nugroho S, Sunyoto J, Lumbanraja H *et al.* Effect of rice variety on methane emission from an Indonesian paddy field. *Soil Sci Plant Nutr* 1997:799–809.

Gogoi N, Baruah KK, Gogoi B *et al.* Methane emission characteristics and its relations with plant and soil parameters under irrigated rice ecosystem of northeast India. *Chemosphere* 2005;**59**:1677–84. https://doi.org/10.1016/j.chemosphere.2004.11.047

Gogoi N, Baruah K, Gupta PK. Selection of rice genotypes for lower methane emission. *Agron Sustain Dev* 2008;**28**:181–6. https://doi.org/10.1051/agro:2008005

Gorh D, Baruah KK. Estimation of methane and nitrous oxide emission from wetland rice paddies with reference to global warming potential. *Environ Sci Pollut Res* 2019;**26**:16331–44.

Gupta K, Kumar R, Baruah KK *et al.* Greenhouse gas emission from rice fields: a review from Indian context. *Environ Sci Pollut Res* 2021;**28**:30551–72. https://doi.org/10.1007/s11356-021-13935-1

Gutierrez J, Atulba SL, Kim G *et al.* Importance of rice root oxidation potential as a regulator of CH4 production under waterlogged conditions. *Biol Fertil Soils* 2014;**50**:861–8. https://doi.org/10.1007/s00374-014-0904-0

Han J, Pei J, Yin Y *et al.* Mining frequent patterns without candidate generation: a frequent-pattern tree approach. *Data Min Knowl Discov* 2004;**8**:53–87. https://doi.org/10.1023/B:DAMI.0000005258.31418.83

Hanson RS, Hanson TE. Methanotrophic bacteria. *Microbiol Rev* 1996;**60**:439–71. https://doi.org/10.1128/mr.60.2.439-471.1996

Huang Y, Sass RL, Fisher FM. Methane emission from Texas rice paddy soils. 1. Quantitative multi-year dependence of CH4 emission on soil, cultivar and grain yield. *Glob Chang Biol* 1997;**3**:479–89. https://doi.org/10.1046/j.1365-2486.1997.00083.x

Husson O. Redox potential (Eh) and pH as drivers of soil/plant/microorganism systems: a transdisciplinary overview pointing to integrative opportunities for agronomy. *Plant Soil* 2013;**362**:389–417. https://doi.org/10.1007/s11104-012-1429-7

Islam MR, Siddique IA, Ali H *et al.* Rice genotypic variation in methane emission patterns under irrigated culture. *Fundam Appl Agric* 2019;**4**:693–703. https://doi.org/10.5455/faa.10569

Jiang Y, Qian H, Wang L *et al.* Limited potential of harvest index improvement to reduce methane emissions from rice paddies. *Glob Chang Biol* 2019;**25**:686–98. https://doi.org/10.1111/gcb.14529

Jiang Y, van Groenigen KJ, Huang S *et al.* Higher yields and lower methane emissions with new rice cultivars. *Glob Chang Biol* 2017;**23**:4728–38. https://doi.org/10.1111/gcb.13737

Jiménez JC, Pedersen O. Mitigation of greenhouse gas emissions from rice via manipulation of key root traits. *Rice* 2023;**16**:24. https://doi.org/10.1186/s12284-023-00638-z

Justin SHFW, Armstrong W. The anatomical characteristics of roots and plant response to soil flooding. *New Phytol* 1987;**106**:465–95. https://doi.org/10.1111/j.1469-8137.1987.tb00153.x

Kim WJ, Bui LT, Chun JB *et al.* Correlation between methane (CH4) emissions and root aerenchyma of rice varieties. *Plant Breed Biotechnol* 2018;**6**:381–90. https://doi.org/10.9787/PBB.2018.6.4.381

Kirk GJD. Rice root properties for internal aeration and efficient nutrient acquisition in submerged soil. *New Phytol* 2003;**159**:185–94. https://doi.org/10.1046/j.1469-8137.2003.00793.x

Kludze HK, DeLaune RD, Patrick WH. Aerenchyma formation and methane and oxygen exchange in rice. *Soil Sci Soc Am J* 1993;**57**:386–91. https://doi.org/10.2136/sssaj1993.03615995005700020017x

Kozela C, Regan S. How plants make tubes. *Trends Plant Sci* 2003;**8**:159–64. https://doi.org/10.1016/S1360-1385(03)00050-5

Lai W-L, Zhang Y, Chen Z-H. Radial oxygen loss, photosynthesis, and nutrient removal of 35 wetland plants. *Ecol Eng* 2012;**39**:24–30. https://doi.org/10.1016/j.ecoleng.2011.11.010

Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;**521**:436–44. https://doi.org/10.1038/nature14539

Li S, Chen L, Han X *et al.* Rice cultivar renewal reduces methane emissions by improving root traits and optimizing photosynthetic carbon allocation. *Agric* 2022;**12**:2134. https://doi.org/10.3390/agriculture12122134

Li Y, Wang X. Root-induced changes in radial oxygen loss, rhizosphere oxygen profile, and nitrification of two rice cultivars in Chinese red soil regions. *Plant Soil* 2013;**365**:115–26. https://doi.org/10.1007/s11104-012-1378-1

Lindau CW, Bollich PK, DeLaune RD. Effect of rice variety on methane emission from Louisiana rice. *Agric Ecosyst Environ* 1995;**54**:109–14. https://doi.org/10.1016/0167-8809(95)00587-I

Lou Y, Inubushi K, Mizuno T *et al.* CH 4 emission with differences in atmospheric CO 2 enrichment and rice cultivars in a Japanese paddy soil. *Glob Chang Biol* 2008;**14**:2678–87. https://doi.org/10.1111/j.1365-2486.2008.01665.x

Luxmoore RJ, Stolzy LH, Letey J. Oxygen diffusion in the soil-plant system IV. Oxygen concentration profiles, respiration rates, and radial oxygen losses predicted for rice roots 1. *Agron J* 1970;**62**:329–32. https://doi.org/10.2134/agronj1970.00021962006200030006x

Ma KE, Qiu Q, Lu Y. Microbial mechanism for rice variety control on methane emission from rice field soil. *Glob Chang Biol* 2010;**16**:3085–95. https://doi.org/10.1111/j.1365-2486.2009.02145.x

Mei XQ, Ye ZH, Wong MH. The relationship of root porosity and radial oxygen loss on arsenic tolerance and uptake in rice grains and straw. *Environ Pollut* 2009;**157**:2550–7. https://doi.org/10.1016/j.envpol.2009.02.037

Niones JM, Suralta RR, Inukai Y *et al.* Field evaluation on functional roles of root plastic responses on dry matter production and grain yield of rice under cycles of transient soil moisture stresses using chromosome segment substitution lines. *Plant Soil* 2012;**359**:107–20.

Noorrohmah S, Takahashi H, Nakazono M. Formation of a barrier to radial oxygen loss in l-type lateral roots of rice. *Plant Root* 2020;**14**:33–41. https://doi.org/10.3117/plantroot.14.33

Philippot L, Andert J, Jones CM *et al.* Importance of denitrifiers lacking the genes encoding the nitrous oxide reductase for N2O emissions from soil. *Glob Chang Biol* 2011;**17**:1497–504. https://doi.org/10.1111/j.1365-2486.2010.02334.x

Qi Z, Guan S, Zhang Z *et al.* Effect and mechanism of root characteristics of different rice varieties on methane emissions. *Agronomy* 2024;**14**:1–15. https://doi.org/10.3390/agronomy14030595

Qin X, Li Y, Wang H *et al.* Effect of rice cultivars on yield-scaled methane emissions in a double rice field in South China. *J Integr Environ Sci* 2015;**12**:47–66. https://doi.org/10.1080/1943815X.2015.1118388

Ranathunge K, Steudle E, Lafitte R. Control of water uptake by rice (*Oryza sativa* L.): role of the outer part of the root. *Planta* 2003;**217**:193–205. https://doi.org/10.1007/s00425-003-0984-9

R Core Team R. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. 2024. https://www.R-project.org/

Revsbech NP, Pedersen O, Reichardt W *et al.* Microsensor analysis of oxygen and ph in the rice rhizosphere under field and laboratory conditions. *Biol Fertil Soils* 1999;**29**:379–85. https://doi.org/10.1007/s003740050568

Sass RL, Fisher FM, Harcombe PA *et al.* Methane production and emission in a TEexas rice field. *Global Biogeochem Cycles* 1990;**4**:47–68. https://doi.org/10.1029/GB004i001p00047

Setyanto P, Rosenani AB, Boer R *et al.* The effect of rice cultivar on methane emission from irrigated rice field. *Indones J Agric Sci* 2004;**5**:20. https://doi.org/10.21082/ijas.v5n1.2004.p20-31

Shin Y-K, Yun S-H. Varietal differences in methane emission from Korean rice cultivars. *Nutr Cycl Agroecosystems* 2000;**58**:315–9. https://doi.org/10.1023/A:1009819324897

Singh S, Kumar S, Jain MC. Methane emission from two Indian soils planted with different rice cultivars. *Biol Fertil Soils* 1997;**25**:285–9. https://doi.org/10.1007/s003740050316

Singh SN, Verma A, Tyagi L. Investigating options for attenuating methane emission from Indian rice fields. *Environ Int* 2003;**29**:547–53. https://doi.org/10.1016/S0160-4120(03)00010-2

Striker GG, Insausti P, Grimoldi AA *et al.* Trade-off between root porosity and mechanical strength in species with different types of aerenchyma. *Plant Cell Environ* 2007;**30**:580–9. https://doi.org/10.1111/j.1365-3040.2007.01639.x

Su J, Hu C, Yan X *et al.* Expression of barley SUSIBA2 transcription factor yields high-starch low-methane rice. *Nature* 2015;**523**:602–6. https://doi.org/10.1038/nature14673

Tamilmounika R, Muthumanickam D, Pazhanivelan S *et al.* Quantifying methane emissions in major rice growing areas of Tamil Nadu using remote sensing and land surface temperature model. *Plant Sci Today* 2024;**11**:1–8. https://doi.org/10.14719/pst.5604

Trolldenier G. Visualisation of oxidizing power of rice roots and of possible participation of bacteria in iron deposition. *Zeitschrift Für Pflanzenernährung Und Bodenkd* 1988;**151**:117–21. https://doi.org/10.1002/jpln.19881510209

Uga Y, Okuno K, Yano M. QTLs underlying natural variation in stele and xylem structures of rice root. *Breed. Sci* 2008;**58**:7–14.

Van Der Welle MEW, Cuppens M, Lamers LPM *et al.* Detoxifying toxicants: interactions between sulfide and iron toxicity in freshwater wetlands. *Environ Toxicol Chem* 2006;**25**:1592–7. https://doi.org/10.1897/05-283R.1

Visser EJW, Bögemann GM. Measurement of porosity in very small samples of plant tissue. *Plant Soil* 2003;**253**:81–90. https://doi.org/10.1023/A:1024560322835

Wang ZP, DeLaune RD, Patrick WH *et al.* Soil redox and pH effects on methane production in a flooded rice soil. *Soil Sci Soc Am J* 1993;**57**:382. https://doi.org/10.2136/sssaj1993.03615995005700020016x

Wang C, Lai DYF, Sardans J *et al.* Factors related with CH4 and N2O emissions from a paddy field: clues for management implications. *PLoS One* 2017;**12**:e0169254. https://doi.org/10.1371/journal.pone.0169254

Wang B, Neue HU, Samonte HP. Role of rice in mediating methane emission. *Plant Soil* 1997;**189**:107–15. https://doi.org/10.1023/A:1004219024281

Wang ZH, Wang LH, Liang H *et al.* Methane and nitrous oxide emission characteristics of high-yielding rice field. *Environ Sci Pollut Res* 2021;**28**:15021–31. https://doi.org/10.1007/s11356-020-11641-y

Wassmann R, Lantin RS, Neue HU *et al.* Characterization of methane emissions from rice fields in Asia. III. Mitigation options and future research needs. *Nutr Cycl Agroecosystems* 2000;**58**:23–36. https://doi.org/10.1023/A:1009874014903

Win KT, Nonaka R, Win AT *et al.* Comparison of methanotrophic bacteria, methane oxidation activity, and methane emission in rice fields fertilized with anaerobically digested slurry between a fodder rice and a normal rice variety. *Paddy Water Environ* 2012;**10**:281–9. https://doi.org/10.1007/s10333-011-0279-x

Xu H, Cai ZC, Tsuruta H. Soil moisture between rice-growing seasons affects methane emission, production, and oxidation. *Soil Sci Soc Am J* 2003;**67**:1147–57. https://doi.org/10.2136/sssaj2003.1147

Yanai RD, Fahey TJ, Miller SL. Efficiency of nutrient acquisition by fine roots and mycorrhizae. In *Physiological Ecology, Resource Physiology of Conifers*, 2013, 73–734.

Yang JX, Liu Y, Ye ZH. Root-induced changes of pH, Eh, Fe(II) and fractions of Pb and Zn in rhizosphere soils of four wetland plants with different radial oxygen losses. *Pedosphere* 2012;**22**:518–27. https://doi.org/10.1016/S1002-0160(12)60036-8

Zhang H, Liu H, Hou D *et al.* The effect of integrative crop management on root growth and methane emission of paddy rice. *Crop J* 2019;**7**:444–57. https://doi.org/10.1016/j.cj.2018.12.011

Zheng H, Fu Z, Zhong J *et al.* Low methane emission in rice cultivars with high radial oxygen loss. *Plant Soil* 2018;**431**:119–28. https://doi.org/10.1007/s11104-018-3747-x