

Integrating traditional knowledge with prompt-based foundation models in rice breeding: perspectives from southern China

Jing-Bo Zhou^{1,2#} , Lei Zhao^{1,2#}, Zhi-Dong Wang^{1,2#}, Qiu-Mei Li^{1,3#}, Rui Jiang⁴, Hong Li^{1,2}, Dao-Qiang Huang^{1,2}, Chong-Rong Wang^{1,2}, Rong Gong^{1,2}, Yang-Yang Pan^{1,2}, Ruo-Wei Miao^{1,2}, Sui-Chun Lai^{1,2}, Yuan-Zhen Ou⁴, Guo-Rong Chen^{1,2*}, Shao-Chuan Zhou^{1,5*} and Yi-Bo Chen^{1,2*}

¹ Rice Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou, Guangdong 510640, China

² Key Laboratory of Genetics and Breeding of High Quality Rice in Southern China (Co-construction by Ministry and Province), Ministry of Agriculture and Rural Affairs, Guangdong Key Laboratory of Rice Science and Technology, Guangzhou, Guangdong 510640, China

³ College of Agriculture, South China Agricultural University and Guangdong Laboratory for Lingnan Modern Agriculture, Guangzhou, Guangdong 510642, China

⁴ College of Engineering, South China Agricultural University and Guangdong Laboratory for Lingnan Modern Agriculture, Guangzhou, Guangdong 510642, China

⁵ Shenzhen Jingmu Meixiang Industrial Co., Ltd., Shenzhen, Guangdong 518000, China

Authors contributed equally: Jing-Bo Zhou, Lei Zhao, Zhi-Dong Wang, Qiu-Mei Li

* Correspondence: chengruorong@gdaas.cn (Chen GR); xxs123@163.com (Zhou SC); chenyibo1213@126.com (Chen YB)

Abstract

Although molecular biology and artificial intelligence are currently prominent research directions in rice (*Oryza sativa*) breeding, the traditional breeding knowledge developed through decades of field practice remains underutilized and infrequently documented. We conducted extensive long-term field investigations in collaboration with experienced breeders in southern China, systematically recording and synthesizing their selection criteria and trait prioritizations to select high-yield and high-quality rice varieties. Agronomic traits that have remained insufficiently characterized in molecular studies were identified, and the strategies breeders apply to integrate these traits to define ideotypes were evaluated. In addition, we assessed the capability of large language models (LLMs) to address breeding inquiries based on practical experience and tested their suitability for applications in the field. Our findings indicated that the current LLMs tend to focus on breeding workflows but do not adequately represent the "art of breeding", including the intuitive inference and prior knowledge that are essential for optimal performance under field conditions. By integrating traditional breeding knowledge with emerging foundation models, this study provides rare documentation of the practices and priorities of traditional breeders, providing new directions for both foundational research and applied breeding, while elucidating how LLMs may be further developed to effectively support agricultural innovation.

Keywords: Rice breeding, Field-based selection, Traditional breeder knowledge, Agronomic ideotypes, Large language models

Citation: Zhou JB, Zhao L, Wang ZD, Li QM, Jiang R, et al. 2026. Integrating traditional knowledge with prompt-based foundation models in rice breeding: perspectives from southern China. *Seed Biology* 5: e026 <https://doi.org/10.48130/seedbio-0026-0022>

Introduction

Rice (*Oryza sativa*) is a major staple crop worldwide and is an important model plant^[1]. Breeding high-yielding and high-quality rice varieties remains the primary objective of crop improvement programs, and elucidating the molecular mechanisms underlying key agronomic traits provides a critical foundation for advancing rice breeding strategies^[2–6]. Over the past 20 years, numerous studies have reported that modifying one or a few genes can notably increase crop yields, with increases ranging from 10% to 68% in crops such as rice and maize (*Zea mays*)^[7–10]. Although these findings have provided valuable insights into molecular biology and gene discovery, most have been evaluated in greenhouse experiments or small-scale field trials, with limited validation in real-world agricultural environments^[11]. For instance, studies have evaluated the effects of 1,671 genes from 47 species on the yield, nitrogen utilization, water efficiency, and other traits of maize. However, following numerous rounds of field trials, only a single gene increased yield by approximately 2%^[11,12]. Major rice research countries such as China and Japan have invested considerable research funds into the exploration and identification of high-yield gene resources^[13–16]. However, despite extensive efforts over the past 20 years, rice yields in these countries have remained stagnant^[17]. These results highlight the considerable disconnect between basic research and practical rice breeding applications.

Although rice has been cultivated for approximately 10 millennia, systematic breeding efforts did not commence until the 1920s. Almost all rice varieties currently cultivated on > 666,700 ha in China were developed using conventional experience-based breeding^[18,19]. Traditional breeders are often overlooked field practitioners, with few records of their practices being preserved. Even without detailed molecular knowledge, they succeeded in breeding breakthrough varieties. A notable example is the semidwarf varieties of the 1960s, which triggered the Green Revolution before the key role of SD1 was identified in 2002^[20,21]. Therefore, a systematic record of traditional breeding practices and agronomic traits, especially those that have received little attention at the molecular level, can provide valuable input for basic research and support the assessment of candidate traits and genes for breeding application.

Attaining expertise as a traditional breeder requires comprehensive agronomic knowledge coupled with substantial field experience, making it challenging for researchers aiming to rapidly transition into applied breeding^[22]. Thus, the efficient integration and extraction of relevant information from extensive datasets are urgently needed. Recent advances in artificial intelligence (AI), particularly large language models (LLMs), present potential tools for aggregating agricultural knowledge and providing comprehensive solutions^[23–26]. Through prompt-based queries, researchers can rapidly acquire specialized knowledge^[27]. However, most of the knowledge accumulated by traditional breeders is transmitted

orally, with limited written documentation, introducing uncertainty regarding the ability of LLMs to accurately capture and reproduce this experiential expertise.

In this study, we systematically compiled and analyzed the selection criteria and ideotypes that traditional rice breeders in southern China have accumulated through decades of experience, many of which have rarely been formally recorded. Subsequently, we evaluated the representation of these same concepts by prompt-based foundation models. By combining breeders' empirically derived heuristics with the models' outputs, we aimed to (1) identify agronomic traits and breeder practices that require further molecular or phenomic investigation, and (2) evaluate the current limitations and potential utility of LLMs for organizing knowledge related to breeding and supporting breeding decisions.

Materials and methods

Field investigations and breeder tracking

We conducted longitudinal field investigations with nine experienced rice breeders at the Rice Research Institute of Guangdong Academy of Agricultural Sciences. The participating breeders were 38–64 years old, all had long-term experience (10 years or more) in field selection of rice, and all worked within southern China's *indica* rice breeding system. In addition, the author visited regions such as Hunan, Hubei, Jiangxi, Guangxi, and Anhui to investigate and collect insights from local markets, farmers, and enterprises. To summarize breeders' knowledge, we focused on traits and decision rules that were repeatedly observed in the field, confirmed by breeders, and supported by examples from local varieties or breeding materials. Field investigations and data collection were conducted over six consecutive growing seasons at several experimental stations from 2022 to 2024, and the observations were synchronized with key crop growth stages. During routine selection activities, we accompanied the breeders in the field, recorded their decision processes in real time, and collected both qualitative descriptions and quantitative measurements of breeders' assessments. Field notes and measurements captured the crop attributes highlighted by breeders during selection, including indicators of the vegetative and reproductive stage such as leaf, panicle, and stem characteristics. Where possible, we supplemented the narrative records with photographs to preserve the visual context for later analysis.

Evaluation of prompt-based foundation models

To assess whether contemporary LLMs can represent breeders' knowledge, we evaluated three publicly available models: SeedLLM-Rice^[24], DeepSeek (V4 pro)^[28], and ChatGPT (GPT-5 Mini). These models were selected to represent two complementary categories: General LLMs that are widely used for scientific information retrieval and knowledge querying, including DeepSeek and ChatGPT, and a foundation model specific to rice, SeedLLM-Rice, which has been developed for rice research and breeding applications. We constructed a small benchmark comprising 20 questions related to breeding experience for model evaluation (Supplementary Tables S1–S20). These questions covered five task categories: (1) Breeding workflow and scope, (2) field selection at the organ level, (3) decision tasks involving multiple traits, (4) direct questions on tacit breeder concepts, and (5) a practical question designed to compare how specialized prompts may improve the practical utility of foundation models in breeding. Model testing was conducted on 2 June 2026. For each model and question, we used the same

prompt template: "You are an experienced rice breeder working in southern China *indica* rice breeding. Answer the following question based on established rice breeding knowledge and practical field-selection logic, with emphasis on practical field observation and selection decisions rather than textbook descriptions alone. If a concept is uncertain, region-specific, or not well documented, explicitly state the uncertainty. Avoid unsupported claims. Provide a concise answer within 300–500 words." Other generation parameters were kept at the platform's default settings. The models' responses were evaluated against breeder knowledge derived from the corresponding evidence described in the manuscript's text. The evaluation framework included six dimensions, each scored from 0 to 2: Accuracy, regional specificity, tacit knowledge coverage, practicality, relevance to field selection decisions, and clarity (Supplementary Table S21). Two experienced rice breeders jointly evaluated the models' responses and reached a consensus on the scores. The scores were used for a structured qualitative comparison of the models' performance. Accuracy evaluated whether the response was consistent with the breeder knowledge supported by the manuscript. Specificity assessed whether the response reflected the ecological and production context of southern China's *indica* rice breeding. Coverage evaluated whether the response captured experience-based traits. Practicality assessed whether the response could guide actual field observations or breeders' decision-making. Relevance assessed whether the response remained focused on field selection rather than shifting to unrelated molecular or statistical descriptions. Clarity assessed whether the answer was sufficiently structured and interpretable for comparison.

Results

Traditional breeders focus on the natural laws of growth and development of rice from sowing to harvest, and they analyze and utilize various traits. For example, early assessment of plant height and leaf and tiller numbers provides a basis for preliminary yield estimation, supporting efforts to enhance grain yield. The interval from booting to maturity is the decisive period for yield and quality formation. This period is characterized by the greatest differences among breeding materials and represents a critical period whereby breeders should focus their efforts on screening. During the rice ripening period, breeders typically perform high-throughput, time-constrained screening, for example, from about 70,000 plants, only approximately 400 of the best strains are selected within 3–7 d, meaning that merely 0.57% of the breeding materials are retained. This process requires the rapid screening of key agronomic traits within a short time. Breeders initially observe traits such as tiller number, ripe leaf color, and panicle curvature and length in the plot population to evaluate the plots with better yield and quality according to their experience. Subsequently, within these selected plots, the yield of a single plant is assessed by hand while shaking the panicles (Supplementary Fig. S1). Traits such as tiller number, stem toughness and thickness, panicle length, grain density, grain length and width, and the seed-setting rate of a single plant were carefully observed and selected. These combined observations established a practical guideline that experienced breeders could use to make fast, simple, and reliable decisions.

Breeders' perspectives on leaf traits

As the primary organs for photosynthesis, the leaves are central to determining rice yield and grain quality. Breeders in southern China share a broad consensus on key leaf traits, including length, width,

number, angle, and growth dynamics. Early in development, sufficient leaf production is required for the accumulation of photosynthetic products. However, excessive numbers or overly large areas of leaves during the later reproductive stage reduce the canopy's permeability and increase susceptibility to diseases. According to breeders' experience, an increased number of leaves does not inherently result in higher yields. At maturity, maintaining only 3–4 functional leaves is considered optimal, as the other leaves, after aging, can redistribute nutrients to the grains, ensuring their full development (Fig. 1a). Breeders describe this plant type as "clean" (Fig. 1b).

Among the last three or four leaves, "ripe color" has long been a defining criterion for high-quality rice in southern China. Breeders select leaves that age slowly, remain physiologically active, and exhibit a slight golden-yellow hue rather than drying up (Fig. 1c, d). Since the 1970s, this phenotype has been an important standard used by breeders to screen and evaluate breeding materials and may involve complex biological processes such as senescence, photosynthesis, and nitrogen metabolism. We observed that this trait reduced the protein content of the grains, thereby improving the eating quality of rice. For instance, 'Meixiangzhan2' has been the most widely promoted variety in Guangdong, China, for the past 9 years. In 1994, breeders used 'Lemont' as parents to generate > 20 hybrid combinations. Its outstanding ripe color has attracted the interest of breeders, who have also determined that its taste quality is excellent. Additionally, certain breeders speculate that the ripe color delays the senescence of leaves, thereby increasing yield. Therefore, ripe color is regarded as a dual indicator of yield potential and grain quality. Notably, although leaf chlorophyll content correlates positively with photosynthetic efficiency, which increases production, some breeders have observed that excessively dark green leaves frequently exhibit poor ripe color, which requires further scientific investigation.

In addition, breeders have reached several consensus points regarding leaf traits. (1) Leaf angle is of practical importance^[29]. During early growth, breeders place minimum focus on the angle (Fig. 1e). However, after the booting stage, a reduced flag leaf angle is preferred to minimize canopy shading and maximize photosynthetic efficiency (Fig. 1f). (2) Leaf length and width segregate substantially. In certain lines, leaves exceeding 1 m in length have been observed (Supplementary Fig. S2). Excessive leaf length is avoided in the humid and typhoon-prone environments in southern China, where long leaves are prone to breakage and bacterial blight. (3) A moderate inward or outward rolling of the leaf is considered to be beneficial, as it can increase structural rigidity (Fig. 1g). (4) Although thicker leaves theoretically enhance photosynthesis, these variants are rare and not widely used in production. Together, these practical criteria form a nuanced framework whereby breeders can assess leaf morphology to balance yield and grain quality.

Breeders' perspectives on panicle traits

Panicle traits are the primary determinants of yield and grain quality. Breeders typically divide rice plants into two types: The multi-panicle type, characterized by numerous panicles, shorter panicles, and fewer grains per panicle in individual plants, and the large-panicle type, characterized by fewer panicles, longer panicles, and more grains per panicle in individual plants, each carrying distinct agronomic trade-offs (Fig. 2a). Breeders have observed that the multi-panicle type fills grains rapidly and achieves a good seed-setting rate, making it suitable across diverse ecological conditions and cultivation systems. Under extreme weather conditions, poor soils, biotic stress, and low-input management, the multi-panicle type confers increased yield stability. For instance, 'HuaZhan' is an important milestone rice variety in China's hybrid rice breeding

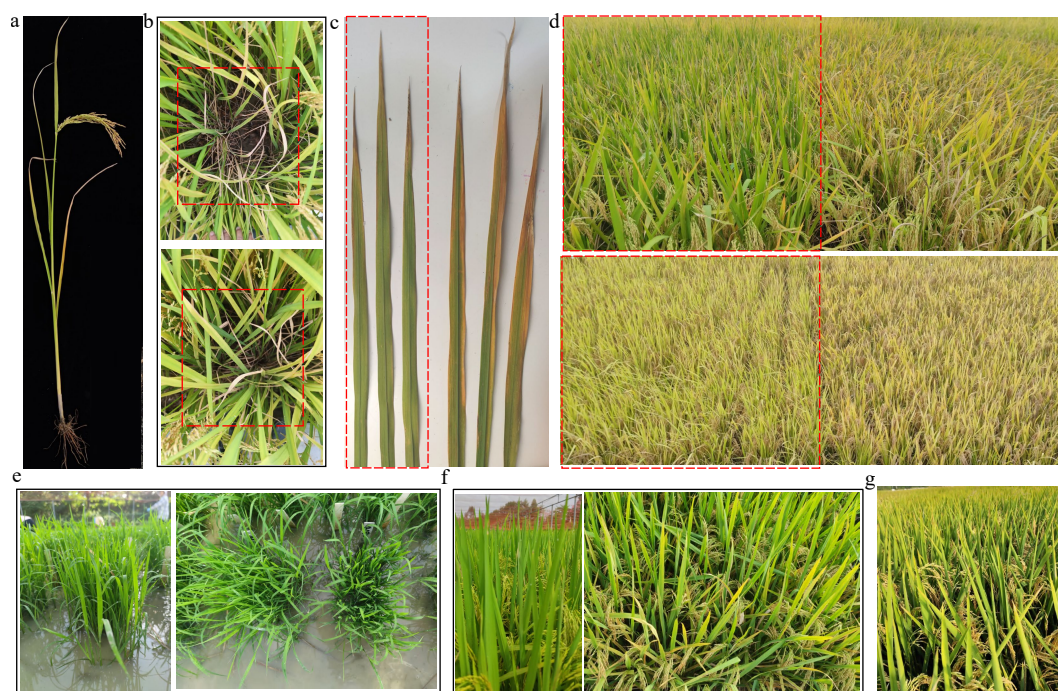


Fig. 1 Perspectives of breeders on leaf traits. (a), (b) Representative "clean" plant type at the maturity stage, characterized by the retention of only three to four functional leaves, shown in (a) a single plant and (b) in the field (clean plant type on the top). (c), (d) Desirable "ripe color" phenotype defined by delayed senescence and a golden-yellow hue, highlighted in the red dotted line boxes in both (c) individual leaves and (d) field views. (e) Leaf angle during early growth, showing the side (left) and top (right) views. (f) Optimal flag leaf angle after booting, showing the side (left) and top (right) views. (g) Slight outward leaf orientation.

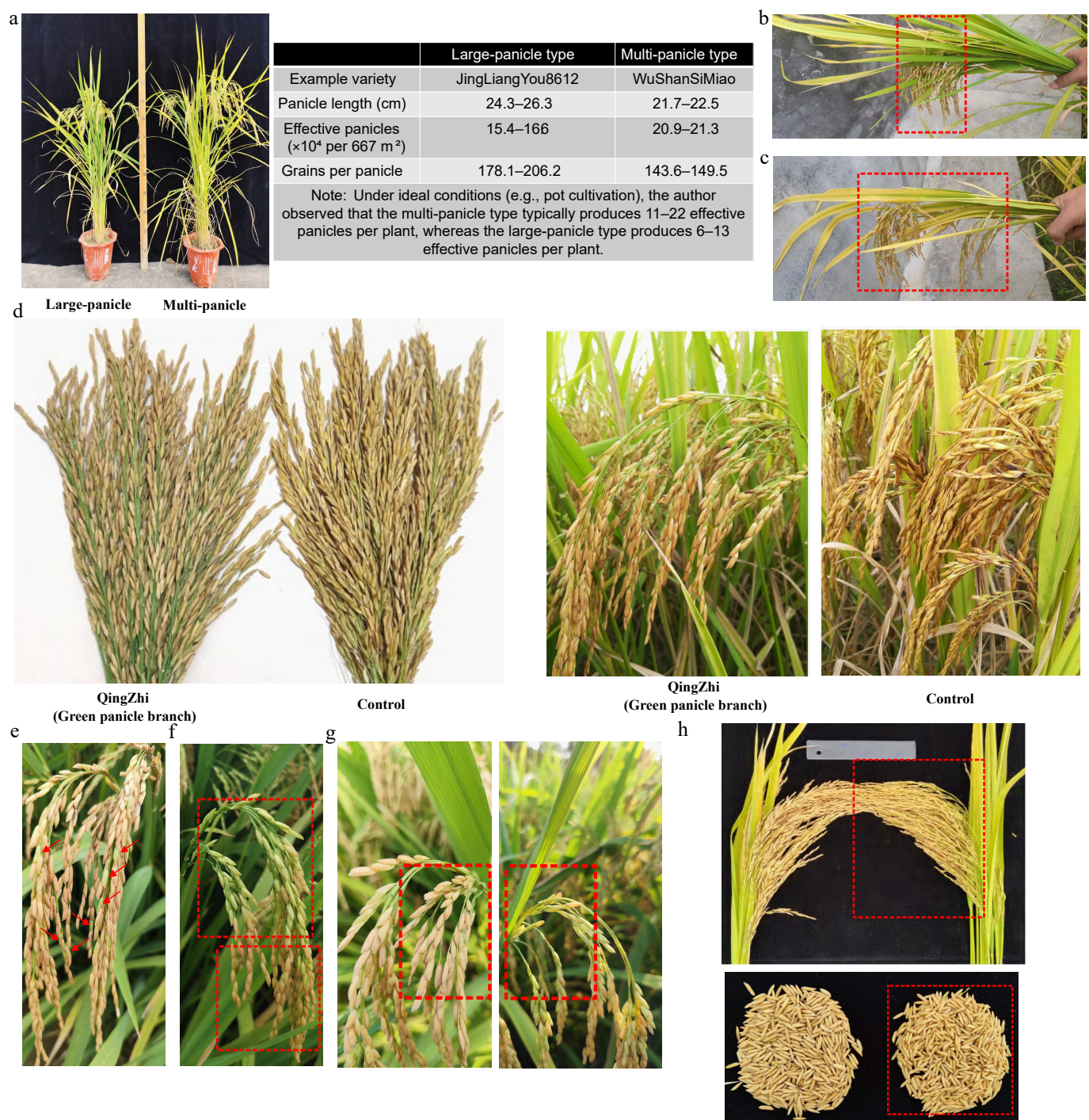


Fig. 2 Breeders' evaluation of panicle-related traits. (a) Illustrations of single plants of the large- and multi-panicle types. The table on the right presents key phenotypic contrasts between representative varieties corresponding to each ideotype, including panicle length, effective panicle number, and grains per panicle. (b) Selection for within-plant uniformity, choosing plants with a similar height and panicle size across tillers. (c) The "big and small panicle" pattern, in which the main tiller carries a larger panicle while the subsequent tillers remain smaller. (d) The green panicle branch (QingZhi) phenotype, defined by the delayed senescence of panicle branches that remain green during the grain-filling stage. Representative field images of this phenotype are shown on the right. (e), (f) Poor seed-set phenotypes: (e) Skipped seeds and (f) two-stage filling that trigger elimination by the breeders. Skipped grains represent the sporadic failure of spikelets to set or fill grains despite normal anthesis. It is characterized by irregularly distributed unfilled spikelets across the panicle (e, red arrows indicate skipped grains). Two-stage grain filling refers to a discontinuous grain development pattern characterized by two temporally distinct stages: The upper spikelets have fully matured, although the basal spikelets have not completed grain-filling (f, red dashed lines). (g) Illustration of poor basal seed set. Red dashed outlines highlight the basal region of the panicle, showing relatively good basal seed set on the left and notably reduced basal fertility on the right. (h) Examples of incomplete glume senescence observed at harvest. Red dashed boxes indicate representative incomplete glumes.

history. As a restorer line, it has been combined with other lines to produce numerous hybrid rice varieties. During the breeding process of 'HuaZhan', breeders identified that the most prominent advantages were its strong tillering ability and large number of panicles. The same traits are also observed in 'WushanSimiao (534)', which serves as the parental line for the hybrid varieties 'JingLiangYou534' and 'LongLiangYou534'. These two hybrids together account for a substantial share of the hybrid rice cultivation area in China. Large-panicle varieties exhibit long panicles with high grain numbers per spike, which are traits that contribute strongly to yield and make these lines visually distinct and preferred by certain regions and breeders. These varieties, such as the typical variety 'JingLiangYou8612', have become popular in middle rice regions such as Anhui and Hunan in China; however, they require increased soil fertility and more intensive management to realize their potential. Both multi-panicle and large-panicle types require further market scientific investigation to determine their actual contributions to yield improvement.

A large diversity of panicle phenotypes appears in the segregating progeny of breeding materials. Breeders predominantly rely on visual observation and tactile evaluation to discern subtle variations in panicle length, 1,000-grain weight, grain density, panicle number, length-to-width ratio, and branch number. Even in bulk populations, experienced breeders can rapidly identify and select the single best plant, whereas lay observers typically perceive no differences. Therefore, key genes or molecular markers are rarely used to screen for panicle traits.

A single plant normally produces multiple tillers, and panicle size typically varies substantially among them. The main tiller has the largest panicle but subsequent tillers show differing panicle sizes, a phenomenon that breeders commonly term "big and small panicles." Using simple assessments holding the rice in their hand, breeders commonly select individual plants wherein panicle sizes across tillers are as similar as possible because this uniformity is believed to support yield stability or even increased yield (Fig. 2b, c, and Supplementary Fig. S1). This trait is strongly influenced by both genetics and environment, and studies on the molecular mechanisms underlying this phenomenon are limited. Breeders believe that when a plant lacks "big and small panicles," the nutritional allocation of the entire plant is balanced, and all the spikes are well developed, which increases yield and quality.

The panicle branch is an important indicator of individual plant yield, frequently evaluated by breeders and basic researchers according to the number of primary and secondary branches. Breeders have summarized a trait known as the "QingZhi" (green panicle branch) phenotype, which refers to branches that exhibit delayed senescence and remain green during the late grain-filling stage (Fig. 2d). Initially reported in 1987, breeders consider this trait to be highly desirable. This trait facilitates the "flow" of nutrients through the branches and stems toward the terminal transport node, ensuring a consistent and robust supply to the grains. This process promotes the development of grains, particularly weak grains, and improves the processing quality of slender grains^[30]. Additionally, this may be beneficial for the breeding of large-panicle rice, as its grain-filling time is long, and weak grains are prone to poor development. "QingZhi" has been promoted as a key feature of conventional and hybrid rice varieties in southern China; however, further investigation is needed to clarify the molecular regulatory mechanisms involved.

The panicle seed-setting rate is an important selection criterion for breeders. Plants with "two-stage grain filling," "skipped grains," or poor basal seed set are eliminated regardless of their other traits (Fig. 2e–g), since seed-setting defects, whether environmental or

genetic, are considered by breeders to be especially difficult to correct. Moreover, the growing frequency of high-temperature events, notably in Hunan, Hubei, and Hefei, has increasingly impaired seed-setting in rice, making this a key breeding target.

Glumes, which protect the floral organs, supply nutrients to the seeds, and determine grain size, have also drawn breeders' attention. Although most glumes have typically senesced by the time breeding materials are selected, certain glumes remain incompletely dried at harvest, a trait that breeders regard as desirable (Fig. 2h). Certain breeders have hypothesized that glume thickness influences heat tolerance. Although an increasing number of studies have investigated the development and photosynthetic activity of glumes, the effects of glume traits on yield, grain quality, and stress resistance require further scientific investigation.

In addition, breeders have pointed out novel directions for improving rice quality. (1) Chalkiness, an appearance quality, has been greatly reduced since the 21st century. However, recent frequent high temperatures have induced chalkiness, severely affecting rice prices. Its frequent occurrence in *indica-japonica* hybrid offspring shows that chalkiness remains a major breeding obstacle. (2) The grain length-to-width ratio is an easily measurable external quality trait that breeders can readily distinguish. Notably, this ratio has shown an increasing trend in southern China, with many newly approved premium varieties currently exceeding 4.0 and several reaching ≥ 4.5 . This trend has even influenced *japonica* rice varieties in Northeast China; however, the increasing length-to-width ratio may compromise processing quality and even hinder yield improvement. (3) Regarding eating quality, most high-quality rice varieties in southern China currently have a medium amylose content. However, eating quality varies significantly among these varieties. Some breeders suggest that modestly reducing the grain protein content while preserving varietal aroma could be an effective strategy for further improvement.

Breeders' perspectives on stem-related traits

The rice stem supports the plant's architecture and mediates nutrient transport, making it essential for an ideal plant type and a key determinant of yield potential. Breeders have consistently focused on lodging resistance, as lodging remains a major constraint to high and stable yields. In 1959, the Guangdong Academy of Agricultural Sciences initiated rice dwarfing programs. Yaoxiang Huang developed 'GuangChangAi', China's first dwarf, high-yield, and lodging-resistant *indica* variety, 7 years earlier than 'IR8' from the International Rice Research Institute, marking a global milestone in rice breeding. Breeders in Guangdong Province typically select lines with thinner, more flexible stems and more tillers on the basis of tactile assessment by hand, as these traits are considered to provide both lodging resistance and yield stability. For instance, during the final selection of 'HuangHuaZhan', 16 individual plants were chosen, and the selected line exhibited reduced plant height, increased stem elasticity, and more tillers. Although provincial trials showed only a modest yield increase (0.5%–3.7%), this variety has strong lodging resistance and adapts well to diverse cultivation systems, including direct-seeded rice, ratoon rice, and rice–crayfish farming. Owing to its broad adaptability to various climates, it is currently the most widely grown *indica* rice variety in China. We also found that although some breeders favor thick stems for lodging resistance, thicker stems do not always ensure better lodging. Currently, breeders frequently use *indica-japonica* hybridization strategies to take advantage of the robust stems of *japonica* materials, a trend that is increasingly evident.

The stem is composed of multiple nodes and internodes. Breeders focus on the neck-spike internode, which refers to the topmost internode on the stems of cereal crops. The length of the internode between the ears is a lodging resistance factor considered by breeders. Breeders suggest that an increased internode length can result in easier lodging (Fig. 3a). Conversely, excessively short internodes may cover the rice spike, subsequently affecting grain development (Supplementary Fig. S3). Breeders also consider internodes in stems as key factors in lodging resistance. When the uppermost nodes do not remain enclosed within the sheath, they may provide weaker support to the panicle, resulting in poor lodging resistance (Fig. 3b). A limited number of breeders also focus on the length of the basal internodes (Fig. 3c, Supplementary Fig. S4), because ground-level lodging presents the greatest yield threat (Supplementary Fig. S5). According to cultivation experts, the first and second internodes at the base notably affect lodging resistance. The plant exhibits strong lodging resistance when the basal internodes are short and thick;

otherwise, the resistance tends to be weaker. Cultivation experts frequently manipulate the basal internode length through the timing of fertilization. Although the relative contributions of the uppermost and basal internodes to lodging resistance require further scientific investigation, many breeders highlight basal internodes as key screening criteria for lodging resistance.

The spatial arrangement of stems among tillers is also an important component of an ideal plant type. Breeders commonly prefer a relatively dispersed arrangement of tiller stems, a feature typical of most modern rice varieties in southern China (Fig. 3d) for the following reasons. First, increased spacing between stems improves light penetration and canopy aeration, enhances leaf photosynthesis, and reduces the incidence of pests and diseases. Second, some breeders believe that dispersed stems may facilitate the dispersal of panicle weight and increase lodging resistance across rice populations (Supplementary Fig. S6). However, these require further scientific investigation.



Fig. 3 Evaluation of stem traits by breeders. (a) Optimal (left) and excessively elongated (right) neck-spike internode. (b) Uppermost node enclosed by the sheath (left, indicated by the red dashed line) and failure of the uppermost node to remain enclosed within the sheath (middle and right, red dashed line). (c) A red dashed box highlights the basal internode length. (d) Dispersed stem arrangement among tillers. A red dashed line indicates the stem orientation. (e) Comparison between typical stems (right) and "lagao" stems (left).

In addition, a "lagao" stem phenotype—observed 30–40 d after flowering and characterized by a waxy yellow color, sustained function, and retained elasticity—has been described by South Chinese breeders since the 1970s (Fig. 3e). To date, breeders still consider this to be a good trait; however, its effects on rice yield and quality require further scientific investigation.

Briefly, as rice yield reaches a bottleneck, lodging resistance is critical for yield breakthroughs, and many high-yield materials are eliminated during breeding because of weak stems. In particular, the ongoing shift toward mechanized, simplified (e.g., direct seeding), large-scale, and diversified cultivation systems has increased the requirements for lodging resistance in rice varieties.

Balancing leaf, panicle, and stem traits

Breeders evaluate panicle, leaf, and stem characteristics together rather than separately to optimize both yield and quality. When presented with a substantial number of breeding materials, breeders initially rely on their experience to select populations (plots) exhibiting superior yield and quality. Subsequently, individual plants are selected through a careful evaluation of comprehensive plant traits, particularly the panicles, leaves, and stems. This pragmatic approach is consistent with Mason and Maskell's source–sink–flow concept, in which the "source" represents photosynthetic organs such as leaves; the "sink" represents organs of consumption or storage, such as panicles and grains; and "flow" represents the vascular pathways, primarily stems, that connect the source and sink. The coordinated function and interactions among source, sink, and flow determine final crop yield. Similarly, traditional breeders infer superior performance by observing multiple organ level traits and their interactions, and subsequently selecting individual plants that best integrate these dynamics.

Two plant types commonly encountered during selection are those with panicles above the flag leaf ("YeShangHe") and those with panicles under the flag leaf ("YeXiaHe"; Fig. 4a, b). Novice selectors tend to favor the YeShangHe type, as their panicle number and size are readily visible. Experienced breeders, however, prefer the YeXiaHe type despite the panicles being partially obscured by the leaves, which makes visual assessment more difficult. Field experience indicates that the latter type more frequently results in higher yield, and this morphology predominates among widely promoted varieties in southern China. Breeders suggest the following possible reasons. First, at maturity, two or three leaves typically remain physiologically active. If the panicle functions in photosynthesis, more

than half of its tissues undergo complete senescence 20–30 d after flowering (Supplementary Fig. S7). Therefore, panicle photosynthesis makes a minor contribution, and its functional duration is relatively short. Consequently, breeders retain erect, active flag leaves at the highest position to sustain the photosynthate supply to the grain, thereby improving both yield and quality. Second, at equal plant heights, the YeXiaHe type exhibits a lower center of gravity, reducing the lodging risk. Third, certain breeders suggest that YeXiaHe has a superior ripe color, although this requires further validation.

In summary, southern China's *indica* plant types cluster into two categories: Multi- and large-panicle types. Multi-panicle types have many effective panicles per plant and smaller panicles per tiller and benefit from stem elasticity and high tiller numbers. Large-panicle types have fewer tillers but larger panicles and therefore require thick stems. For both ideotypes, breeders commonly selected erect flag leaves, good ripe color, uniform panicle size across tillers, high seed-setting rates, and green panicle branches. Simultaneously, the right types of panicles under the flag leaf are required. Breeders also use stem internode traits to improve lodging resistance. In terms of grain quality, breeders prioritize low chalkiness, high transparency, fragrance, and an increased length-to-width ratio. Collectively, these field-based selection criteria not only define the plant types preferred by breeders but also highlight traits that require further investigation. This prioritization reflects breeders' repeated observations, field relevance, existing evidence, and remaining research gaps. Some trait functions are supported by scientific evidence but require deeper genetic characterization and breeding evaluation, whereas other proposed effects remain to be rigorously validated (Supplementary Table S22). Together, these criteria guide breeders in balancing the yield potential, quality, and resistance of South Chinese rice varieties.

Performance evaluation of LLMs in breeding experience

To evaluate whether foundation models guided by prompts can represent breeding experience rather than merely reproduce textbook-based breeding knowledge, we constructed a small benchmark comprising 20 questions (Fig. 5, Supplementary Tables S1–S20). These questions were designed to assess whether ChatGPT, DeepSeek, and SeedLLM-Rice could respond to breeding tasks in a manner consistent with the field practices and decision-making of experienced breeders in southern China. The models' responses



Fig. 4 Summary of leaf, panicle, and stem traits. (a) Comparative plant architectures illustrating panicles positioned above the flag leaf (YeShangHe) and below the flag leaf (YeXiaHe). (b) YeShangHe and YeXiaHe under field conditions.

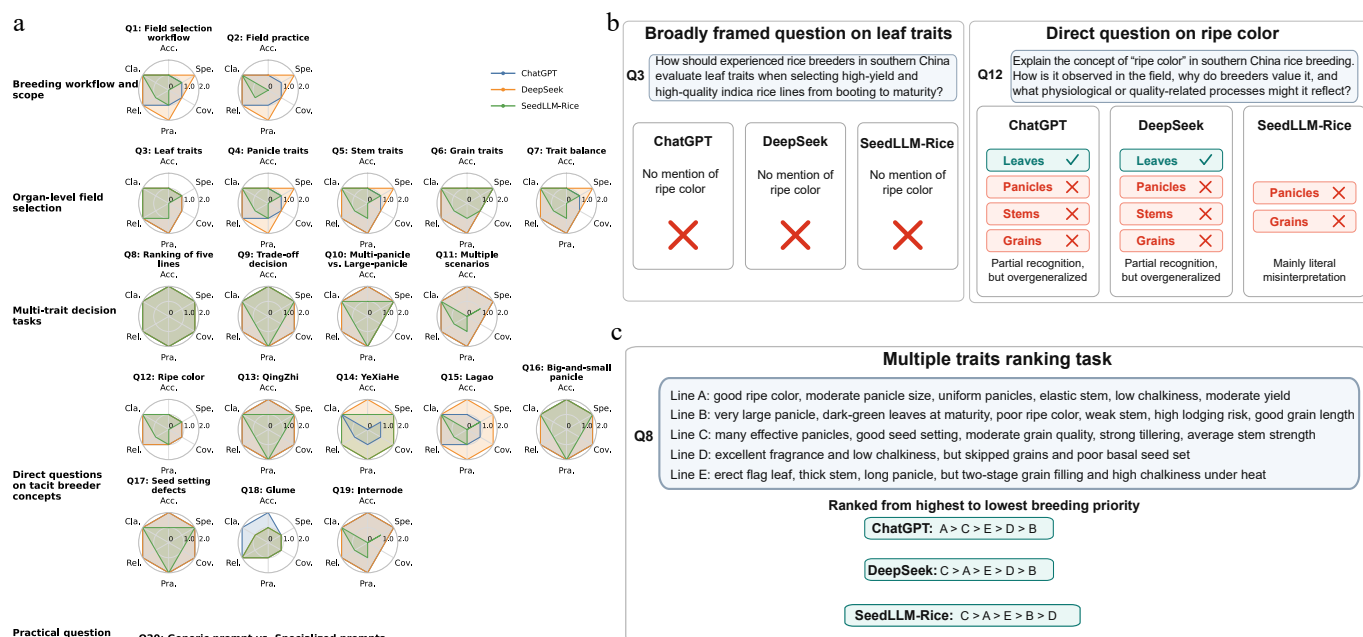


Fig. 5 Comparative scoring and representative responses of LLMs across rice breeding questions. (a) Radar plots summarizing the performance of ChatGPT, DeepSeek, and SeedLLM-Rice across rice breeding questions. Each radar plot includes six scoring criteria, namely accuracy (Acc.), clarity (Cla.), relevance (Rel.), practicality (Pra.), coverage (Cov.), and specificity (Spe.), with scores ranging from 0 to 2. Questions were grouped into five categories. (b) Representative LLM responses to a broadly framed question and a direct question on tacit breeder knowledge. (c) Representative LLM responses to a ranking task involving multiple traits.

were evaluated using six assessment criteria to examine their alignment with the breeder knowledge described above. DeepSeek achieved the highest score, with ChatGPT showing comparable overall performance (Fig. 5a). We therefore further examined the content of the models' responses to identify how the three models differed in their representation of breeder knowledge. For questions related to breeding workflow and scope (Supplementary Tables S1–S2), all three models were able to describe the general logic of conventional pedigree breeding. DeepSeek showed stronger alignment with the practical scale of breeding operations, although it also introduced several unsupported numerical or operational details. ChatGPT provided a clear and generally relevant description of the workflow, but its response remained more conceptual than fully applicable to field practice. SeedLLM-Rice described a structured workflow but provided more limited coverage of tacit field traits.

The contrast between a broadly framed question on leaf traits and a direct question on ripe color further revealed the difficulty of eliciting tacit breeder knowledge. Though the question broadly asked how experienced breeders evaluate leaf traits without explicitly mentioning ripe color (Fig. 5b, Supplementary Table S3), none of the three models fully reproduced the central role of ripe color as a tacit selection criterion in southern China's rice breeding. When ripe color was queried directly (Fig. 5b, Supplementary Table S12), both ChatGPT and DeepSeek recognized it as a traditional field trait. However, their responses expanded the concept beyond functional leaves to include the overall visual appearance of panicles, stems, or grains. SeedLLM-Rice showed the strongest misinterpretation, mainly explaining ripe color as panicle or grain color. These results suggest that broadly framed prompts may fail to elicit tacit breeder knowledge, whereas direct questioning can recover only part of this knowledge. Although explicit prompting enabled the models to partially recognize the concept, their understanding remained inconsistent, often shifting toward overly general or literal interpretations.

To further assess integrated breeding judgment, we used a ranking task involving multiple traits, which provided the clearest evidence that the models could approximate integrated breeding decisions when key traits and defects were explicitly specified (Fig. 5c, Supplementary Table S8). All three models recognized the need to balance yield potential, grain quality, lodging resistance, ripe color, and seed-setting, rather than ranking materials according to a single attractive trait. Their responses generally reflected the practical breeding principle that major defects can outweigh individual advantages. These results indicate that in structured tasks with explicit trait conflicts, the models can approximate the reasoning used by breeders to integrate multiple traits. We next evaluated whether the models could identify negative traits that would lead breeders to eliminate a line, rather than treating all traits as neutral descriptors (Supplementary Table S17). ChatGPT and DeepSeek both aligned closely with the elimination logic. SeedLLM-Rice also recognized these traits as undesirable but introduced gene- or omics-related explanations that were not central to the logic of field selection. This result suggests that the models performed better when traits associated with negative selection were explicitly specified, whereas SeedLLM-Rice was more likely to shift toward molecular explanations.

Finally, to illustrate how specialized prompts may improve the practical utility of foundation models in breeding, we designed two semantically related prompts on phenomic workflows for quantifying ripe color (Supplementary Table S20). Under the generic prompt, ChatGPT and DeepSeek generated broad but relatively balanced phenomics workflows. When the prompt explicitly defined ripe color in the context of southern China's rice breeding, ChatGPT and DeepSeek generated more complete workflows with stronger relevance to field practice. Under the specialized prompt, SeedLLM-Rice improved from a grain-focused answer generated under the generic prompt to a more appropriate framework based on leaf senescence, although its workflow remained simpler and less detailed. These

results suggest that specialized prompt design can help convert tacit breeder knowledge into more implementable phenomic evaluation workflows.

Overall, ChatGPT tended to provide a clearer structure and more organized decision logic, DeepSeek tended to provide more detailed and operationally framed responses, and SeedLLM-Rice tended to generate rice-relevant responses, but these were more strongly focused on molecular explanations. Current foundation models, guided by prompts, can help summarize general knowledge on rice breeding and generate preliminary checklists for field selection. However, their outputs should be validated against breeder knowledge before being used to support breeding decisions. Therefore, these models may be useful as tools for organizing knowledge, but they do not yet fully capture the "art of breeding," particularly the tacit field judgments used by experienced rice breeders in southern China.

Discussion

Rice is a globally important staple, and developing high-yield, high-quality varieties is key to food security. Breeders have always been central to this effort, and understanding their methods supports the exploration of gene function, technology innovation, and breeding efficiency gains. Although recent omics advances, including genomics, transcriptomics, and proteomics, have uncovered many genes related to agronomic traits, traditional breeders still hold clear advantages in practical variety development^[15,16].

First, with respect to agronomic traits, breeders emphasize multiple traits across the entire rice growth cycle, especially from grain filling to maturity, rather than focusing on individual traits or genes, with the goal of improving yield and quality^[5]. Achieving an optimal trait balance is a primary objective. For instance, although 1,000-grain panicle materials were developed in the 1960s, they failed to translate into significant yield gains under field conditions. Moreover, they often make extensive use of favorable agronomic traits even when the underlying genes and mechanisms remain unclear.

Second, rather than being grown in pots or greenhouses, breeding materials are planted under high-density field conditions with a narrow row spacing, thereby subjecting them to the full complexity of natural environmental stresses. Breeders focus on the population rather than on individual plants, as individual yield does not correlate strongly with population yield. Additionally, they not only prioritize yield and quality under a given condition but also emphasize stable performance across a range of ecological environments. The timing and frequency of irrigation and fertilizer application are strictly aligned with actual production requirements. The materials then undergo 2–3 additional years of multi-location field trials across 14–20 ecological sites. Most approved cultivars show a yield increase of 3%–5%, with only a few reaching as high as 10%. Even after approval, however, only a small fraction are successfully promoted on a large scale.

Third, breeders maintain a distinct advantage in the continuous improvement of the agronomic traits of rice, as varieties older than 15–20 years are rarely retained because of their declining performance in terms of resistance, yield, and quality. Instead, they advance breeding by pyramiding superior genes and traits into current elite cultivars^[31]. Furthermore, breeders account for future market demand and the potential for large-scale dissemination when developing new varieties, thereby exhibiting foresight, adaptability, and innovation. For instance, in regions where shorter growth periods are favored to mitigate extreme weather risks and facilitate crop

rotations, breeders have responded with targeted adaptations. This is an important factor underlying the superior yield, quality, and resistance of modern rice varieties.

Additionally, rice is a model plant, with '9311' and 'Nipponbare' (released in the 1990s and 1950s, respectively) serving as reference materials for validating genetic effects on agronomic traits. Yet these old varieties show considerable gaps in yield, quality, and resistance relative to the currently dominant cultivars. As a result, transgenic validation data, especially for yield and some quality traits, often do not accurately predict the practical utility of target genes^[11,12,17]. Furthermore, genes cloned from such outdated materials may have already been deployed, discarded, or superseded by superior alleles in elite breeding lines. Therefore, gene validation using widely grown varieties across multiple ecological sites should offer a more thorough evaluation of practical genetic value.

Breeders currently face numerous challenges. First, although breeders have been effective in breeding materials for yield and quality, changing climatic conditions and new market demands have created additional requirements, such as tolerance to flooding, high temperature, low temperature, low cadmium, and herbicides, as well as adaptation to diverse cultivation systems including direct seeding and mechanization. However, for many of these traits, relevant germplasm resources and genes remain insufficient. Second, despite the identification and development of numerous key genes as molecular markers, gene pleiotropy constrains the number of markers applicable in practice, resulting in limited uptake by breeders. They therefore continue to depend on long-standing personal expertise, which inevitably introduces some uncertainty. Third, the narrowing genetic base of rice breeding materials increasingly limits further improvements in yield and quality. Expanding genetic diversity, especially via *indica-japonica* hybridization, is therefore a key breeding strategy^[32]. Over decades of indirect introgression, many current varieties have come to carry partial ancestry from both subspecies. Our observations indicate that seed-setting in *indica-japonica* hybrids is no longer a major constraint; instead, the main challenges involve trade-offs among yield, quality, and resistance. Although some breeders advocate for prioritizing yield while enhancing quality and resistance, the development of novel methods and resources, along with sustained effort, remains essential.

Prompt-based foundation models provide apparent benefits for knowledge dissemination^[25,27]. With well constructed prompts, trainees can quickly learn breeding theories and routine methods without an exhaustive literature review. However, foundation models encounter challenges in addressing tacit, sensory, and context-dependent knowledge that breeders accumulate over the years of fieldwork. In our comparisons, the models handled the conventional theory well. However, they produced generic responses on selection details, and certain outputs overemphasized analytical pipelines such as genome-wide association studies and statistical genetics that are peripheral to frontline phenotypic selection. Two possible explanations are that much breeder knowledge is transmitted orally and is thus underrepresented in training corpora, and that current prompts are insufficient to elicit detailed, practice-oriented guidance. To address these limitations, at least three approaches should be pursued. First, owing to the diversity of crop species and substantial variation in climate and soil conditions across regions, universal models are unlikely to be completely effective^[24]. Dedicated region-specific models and specialized modules are necessary. Second, the current application of foundation models in agriculture remains confined mainly to small-scale trials within research settings, with limited implementation in public services. To extend their utility to a broader range of users, including traditional breeders and farmers, models must provide greater interpretability and transparency

so that users can understand and trust their output^[25,27]. Finally, the development of more specialized and comprehensive prompt libraries can improve the accuracy of model responses while reducing entry barriers for the end users.

Future work should integrate basic research with practice, exploit key genes of proven value, and use phenomics to quantify breeders' experience, thereby enhancing efficiency. Prompt-based models can support knowledge dissemination but must be trained on curated, practice relevant data and validated by breeders before field application.

Conclusions

This study systematically documents practical, experience based selection criteria used by traditional rice breeders in southern China and evaluates how these tacit breeding concepts are represented by prompt-based foundation models. Through long-term field investigations and breeder tracking, we identified a set of traits prioritized by breeders. Although these traits are routinely used by experienced breeders to balance yield potential, grain quality, lodging resistance, and field adaptability, many remain insufficiently characterized in molecular and phenomic studies.

Ethical statements

This study was conducted in accordance with the Declaration of Helsinki. Institutional review board approval was waived by the Ethics Committee of Guangdong Academy of Agricultural Sciences Rice Research Institute because the research involved only nonsensitive, professional interviews with experts.

Verbal informed consent was obtained from all participants prior to the interviews. Verbal consent was used instead of written consent because the study focused solely on professional breeding experience, involved no risk or invasive procedures, and collected no sensitive personal or health-related data. All data were anonymized to ensure the participants' confidentiality.

Author contributions

The authors confirm their contributions to the paper as follows: study conception and design: Zhou JB, Zhou SC, Chen GR, Chen YB; field investigation and data collection: Zhou JB, Zhao L, Wang ZD, Li QM, Li H, Huang DQ, Wang CR, Gong R, Pan YY, Miao RW, Lai SC, Ou YZ, Zhou SC, Chen YB; analysis and interpretation of the results: Zhou JB, Zhao L, Wang ZD, Li QM, Jiang R, Zhou SC, Chen GR, Chen YB; figure preparation, visualization and draft manuscript preparation: Zhou JB, Zhao L, Wang ZD, Li QM; manuscript revision and critical editing: Zhou SC, Chen GR, Chen YB. All authors reviewed the results and approved the final version of the manuscript.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Acknowledgments

We gratefully acknowledge the Chinese scientists, breeders, and farmers whose exceptional contributions and dedication have significantly advanced rice breeding in China. Their expertise has been indispensable in driving agricultural innovation and progress. This work was supported by National Key Research and Development

Program (No. 2024YFD1200801), the National Natural Science Foundation of China (NSFC) (No. 32301707), the Joint Funds of the National Natural Science Foundation of China (No. U22A20464), the Key R&D Special Project of Henan Province (No. 26111112600); the Shenzhen Science and Technology R&D Fund (No. JSGGKQTD20221101115652015), the Special Project for Science and Technology Innovation Strategy (No. R2023PY-JX002), the Guangzhou Key Research and Development Program (Nos. 2024B03J1319, 2024B03J1356), the Guangdong Key Laboratory of Rice Science and Technology (No. 2023B1212060042), the Guangdong Association for Science and Technology Young Sci-Tech Talent Cultivation Program (No. SKXRC2025532), Natural Science Foundation of Guangdong Province (No. 2024A1515012608), the Guangdong Province Rural Revitalization Strategy Special Fund Seed Industry Revitalization Project (No. 2024-NPY-00-009), the Science and Technology Talent Introduction Special Fund Project of Guangdong Academy of Agricultural Sciences (No. R2024YJ-YB3007), Special Funding for the Construction of the High-Level Academy of Agricultural Sciences (NYQS202630), and the Elite Rice Plan of GDRRI (Nos.2023YG03; 2025YG05).

Conflict of interest

Zhou SC is affiliated with Shenzhen Jingu Meixiang Industrial Co., Ltd. The company did not provide any funding or patented materials. All authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary information accompanies this paper online at: <https://doi.org/10.48130/seedbio-0026-0022>.

Dates

Received 6 May 2026; Revised 5 August 2026; Accepted 3 September 2026; Published online 14 October 2026

References

- [1] Chen R, Deng Y, Ding Y, Guo J, Qiu J, et al. 2022. Rice functional genomics: decades' efforts and roads ahead. *Science China Life Sciences* 65:33
- [2] Huang J, Li J, Zhou J, Wang L, Yang S, et al. 2018. Identifying a large number of high-yield genes in rice by pedigree analysis, whole-genome sequencing, and CRISPR-Cas9 gene knockout. *Proceedings of the National Academy of Sciences of the United States of America* 115:E7559–E7567
- [3] Ren D, Ding C, Qian Q. 2023. Molecular bases of rice grain size and quality for optimized productivity. *Science Bulletin* 68:314–350
- [4] Li P, Chen YH, Lu J, Zhang CQ, Liu QQ, et al. 2022. Genes and their molecular functions determining seed structure, components, and quality of rice. *Rice* 15:18
- [5] Li G, Tang J, Zheng J, Chu C. 2021. Exploration of rice yield potential: decoding agronomic and physiological traits. *The Crop Journal* 9:577–589
- [6] Liu C, Wang J, Lu H, Huang Y, Yan H, et al. 2024. Engineering synthetic apomixis in different hybrid rice varieties using the *Fix* strategy. *New Crops* 1:100003
- [7] de Souza AP, Burgess SJ, Doran L, Hansen J, Manukyan L, et al. 2022. Soybean photosynthesis and crop yield are improved by accelerating recovery from photoprotection. *Science* 377:851–854
- [8] Kromdijk J, Glowacka K, Leonelli L, Gabilly ST, Iwai M, et al. 2016. Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. *Science* 354:857–861

- [9] Wei S, Li X, Lu Z, Zhang H, Ye X, et al. 2022. A transcriptional regulator that boosts grain yields and shortens the growth duration of rice. *Science* 377:eabi8455
- [10] Yu Q, Liu S, Yu L, Xiao Y, Zhang S, et al. 2021. RNA demethylation increases the yield and biomass of rice and potato plants in field trials. *Nature Biotechnology* 39:1581–1588
- [11] Khaipho-Burch M, Cooper M, Crossa J, de Leon N, Holland J, et al. 2023. Genetic modification can improve crop yields—but stop overselling it. *Nature* 621:470–473
- [12] Simmons CR, Lafitte HR, Reimann KS, Brugière N, Roesler K, et al. 2021. Successes and insights of an industry biotech program to enhance maize agronomic traits. *Plant Science* 307:110899
- [13] Chen H, Xie W, He H, Yu H, Chen W, et al. 2014. A high-density SNP genotyping array for rice biology and molecular breeding. *Molecular Plant* 7:541–553
- [14] Xu X, Liu X, Ge S, Jensen JD, Hu F, et al. 2012. Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nature Biotechnology* 30:105–111
- [15] Wing RA, Purugganan MD, Zhang Q. 2018. The rice genome revolution: from an ancient grain to Green Super Rice. *Nature Reviews Genetics* 19:505–517
- [16] Miura K, Ashikari M, Matsuoka M. 2011. The role of QTLs in the breeding of high-yielding rice. *Trends in Plant Science* 16:319–326
- [17] Xiong D, Flexas J, Huang J, Cui K, Wang F, et al. 2022. Why high yield QTLs failed in preventing yield stagnation in rice? *Crop and Environment* 1:103–107
- [18] Zheng W, Ma Z, Zhao M, Xiao M, Zhao J, et al. 2020. Research and development strategies for hybrid japonica rice. *Rice* 13:36
- [19] Cheng SH, Zhuang JY, Fan YY, Du JH, Cao LY. 2007. Progress in research and development on hybrid rice: a super-domesticated in China. *Annals of Botany* 100:959–966
- [20] Spielmeier W, Ellis MH, Chandler PM. 2002. Semidwarf (*sd-1*), "green revolution" rice, contains a defective gibberellin 20-oxidase gene. *Proceedings of the National Academy of Sciences of the United States of America* 99:9043–9048
- [21] Sasaki A, Ashikari M, Ueguchi-Tanaka M, Itoh H, Nishimura A, et al. 2002. A mutant gibberellin-synthesis gene in rice. *Nature* 416:701–702
- [22] Repinski SL, Hayes KN, Miller JK, Trexler CJ, Bliss FA. 2011. Plant breeding graduate education: opinions about critical knowledge, experience, and skill requirements from public and private stakeholders worldwide. *Crop Science* 51:2325–2336
- [23] Zhang R, Wang Y, Yang W, Wen J, Liu W, et al. 2025. PlantGPT: an Arabidopsis-based intelligent agent that answers questions about plant functional genomics. *Advanced Science* 12:e03926
- [24] Yang F, Kong H, Ying J, Chen Z, Luo T, et al. 2025. SeedLLM-Rice: a large language model integrated with rice biological knowledge graph. *Molecular Plant* 18:1118–1129
- [25] Yin S, Xi Y, Zhang X, Sun C, Mao Q. 2025. Foundation models in agriculture: a comprehensive review. *Agriculture* 15:847
- [26] Zhu W, Han R, Shang X, Zhou T, Liang C, et al. 2024. The CropGPT project: call for a global, coordinated effort in precision design breeding driven by AI using biological big data. *Molecular Plant* 17:215–218
- [27] Tzachor A, Devare M, Richards C, Pypers P, Ghosh A, et al. 2023. Large language models and agricultural extension services. *Nature Food* 4:941–948
- [28] DeepSeek-AI, Xu A, Lin B, Xue B, Wang B, et al. 2026. DeepSeek-v4: towards highly efficient million-token context intelligence. *arXiv Preprint:2606.19348*
- [29] Cui W, Ji C, Xu Y, Zhou H, Hong L. 2026. Spatiotemporal regulation of leaf angle in rice: the differential regulation of leaf angle throughout the development and across the canopy in rice. *New Crops* 3:100088
- [30] Chen Y, Zhao L, Wang C, Li H, Huang D, et al. 2023. Stay-green panicle branches improve processing quality of elite rice cultivars. *Rice Science* 30:11–14
- [31] Zheng X, Wei F, Cheng C, Qian Q. 2024. A historical review of hybrid rice breeding. *Journal of Integrative Plant Biology* 66:532–545
- [32] Zhang G. 2022. The next generation of rice: inter-subspecific indica-japonica hybrid rice. *Frontiers in Plant Science* 13:857896



Copyright: © 2026 by the author(s). Published by Maximum Academic Press on behalf of Hainan Yazhou Bay Seed Laboratory. This article is an open access article distributed under Creative Commons Attribution License (CC BY 4.0), visit <https://creativecommons.org/licenses/by/4.0/>.