



# Induction and characterization of a novel high-protein mutant in peanut (*Arachis Hypogaea* L.)

Shu Tao Yu<sup>1</sup> · Xiang Yin<sup>2</sup> · Ye Chao Yin<sup>1</sup> · Yu Zhang<sup>1</sup> · Jing Chao Dong<sup>1</sup> · Lei Lei Gao<sup>1</sup> · Hao Jie Sun<sup>2</sup> · Guang Di Yuan<sup>2</sup> · Chun Jiao Jiang<sup>2</sup> · Bei Bei Li<sup>2</sup> · Chuan Tang Wang<sup>2</sup> · Jing Yu<sup>2</sup>

Received: 11 March 2025 / Accepted: 24 January 2026

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2026

## Abstract

Peanut, a globally significant food crop, is highly valued for its nutritional and economic importance. However, breeding efforts to enhance its protein content have been limited. In this study, a two-step chemical mutagenesis strategy was utilized on the normal-protein Spanish-type peanut cultivar Fuhua 19 (<25% protein) to generate high-protein mutants. Through successive rounds of mutagenesis combined with near-infrared spectroscopy screening, eight mutants with protein levels exceeding 30% were identified. Among these, the mutant C-Za-454-2 consistently exhibited stable and elevated protein content across different locations and developmental stages. This mutant also displayed distinct ultrastructural changes, including an increased protein body area and reduced starch grain content, compared to the wild-type Fuhua 19. Biochemical analyses revealed enhanced activities of nitrogen metabolism enzymes, such as nitrate reductase and glutamate synthetase, which positively correlated with protein accumulation. Conversely, certain carbon metabolism enzyme activities were inversely correlated with protein content, highlighting a trade-off between carbohydrate and protein synthesis. These findings demonstrate the efficacy of iterative mutagenesis in producing high-protein peanut lines and provide a foundation for future genetic studies aimed at improving protein yield in peanut cultivars.

**Keywords** Groundnut · Protein content · Ethyl methanesulfonate · Near-infrared spectroscopy · Protein metabolism · Ultrastructure analysis · Nitrogen assimilation

Communicated by Mohsin Tanveer.

✉ Chuan Tang Wang  
chinapeanut@126.com

✉ Jing Yu  
iamyujing2008@126.com

Shu Tao Yu  
shutaoyu@163.com

Ye Chao Yin  
yechao202411@163.com

Yu Zhang  
zhangyu611426@163.com

Jing Chao Dong  
djchlt@163.com

Lei Lei Gao  
gao1029815131@outlook.com

Hao Jie Sun  
1191834745@qq.com

Guang Di Yuan  
yuangd42@163.com

Chun Jiao Jiang  
qaujcj@163.com

Bei Bei Li  
825859495@qq.com

<sup>1</sup> Institute of Sandy Land Management and Utilization of Liaoning, Fuxin 123000, China

<sup>2</sup> Shandong Peanut Research Institute, Qingdao 266100, China

## Introduction

The cultivated peanut (*Arachis hypogaea* L.) is a globally important cash crop prized for its nutritional and economic value (Woodroof 1983; Stalker and Wilson 2016; Smartt 1994; Mallikarjuna and Varshney 2014; Lusas et al. 1989; Nwokolo and Smartt 1996). Renowned for its high oil content, highly digestible protein, and abundant health-promoting phytochemicals, peanut plays a crucial role in combating malnutrition and bolstering food security, particularly in developing and low-income regions (Whitworth 2021; Chen et al. 2024; Sakha et al. 2022). Among its agronomic traits, protein content is a primary focus for improvement due to its importance in meeting dietary protein needs and its applications in value-added food products. Enhancing protein content in peanut is not only pivotal for improving nutritional quality but also for expanding the economic viability of peanut-derived industries (Zhao et al. 2023; Singh and Singh 1991; Varshney et al. 2017). Despite the  $\geq 28\%$  protein content requirement for cultivars for food uses mandated by China's National Peanut Regional Trials and substantial market demand from peanut butter manufacturers for high-protein raw materials, the development of high-protein peanut varieties has progressed significantly slower compared to breeding achievements in high-oleic acid ( $\geq 75\%$ ) and high-oil ( $\geq 55\%$ ) traits (Vargheese et al. 2024; Janila et al. 2016; Yang et al. 2023; Huang et al. 2023).

Being the product of evolution after a series of six bottlenecks, the cultivated peanut is an amphidiploid species with a narrow gene base (Mallikarjuna and Varshney 2014). High-protein germplasm resources in peanut are largely concentrated in small-seeded genotypes including the Valencia-type cultivars, which is characterized by lower yields. While some of the wild *Arachis* species have been reported to possess high-protein trait (Li et al. 1994), the use of these resources is often hindered by linkage drags, whereby undesirable traits are co-introduced along with the desired ones during hybridization (Stalker 2017; Kassie et al. 2023). Consequently, exploiting wild species for breeding high-protein cultivars remains a challenging and labor-intensive process. This underscores the need for alternative approaches to broaden the gene base for high-protein peanut breeding. One such approach is the use of mutagenesis to induce variation in high-performing cultivars with superior agronomic traits, thereby overcoming the limitations associated with wild species and low-yielding germplasm.

Considering the amphidiploid nature of the cultivated peanut, our team proposed a two-step mutagenesis strategy (Yu et al. 2019), involving successive rounds of mutagenesis to enhance genetic variability and increase the likelihood of achieving desirable phenotypic changes. This approach has demonstrated success in the development of high-oleic

peanut varieties (Yu et al. 2019). However, its applicability to breeding high-protein peanut mutants remains unexplored. Unlike high-oil or high-oleic variants, high-protein peanut mutants have been rarely documented and lack comprehensive characterization (Yu et al. 2015; Liang et al. 2022).

The objectives of this study were to (1) evaluate the effectiveness of two rounds of mutagenesis in generating high-protein peanut mutants, (2) assess the stability of protein content in a high-protein mutant across two distinct locations, (3) elucidate the ultrastructural differences in cotyledon tissues between the high-protein mutant and the wild-type peanut, and (4) investigate the differences in nitrogen and carbon metabolism enzyme activities, protective enzyme activities, and malondialdehyde (MDA) levels between the high-protein mutant and wild-type peanut, with an emphasis on exploring the relationships between these biochemical indicators and protein content.

## Materials and methods

### Peanut materials and cultivation

The normal-protein Spanish-type peanut variety Fuhua 19 (FH19), developed at the Liaoning Sandy Land Improvement and Utilization Research Institute (ISLMUL), along with its mutagenized populations, was cultivated at the ISLMUL Breeding Base (121°70'E, 42°06'N), High-Tech Zone, Liaoning Province, China.

The high-protein mutant C-Za-454-2, initially identified at the Breeding Base, was grown at the Sihewa Base (121°68'E, 42°02'N), Wangfu Town, Fuxin County, Liaoning Province, for preliminary testing.

The estimation of chlorophyll content and sampling and measurement of 14 biochemical indicators at different growth stages for the high-protein mutant C-Za-454-2 and wild-type FH9 were performed at the Breeding Base. To calculate the correlation coefficients between biochemical indicators at different stages and protein content at maturity, in addition to the high-protein mutant C-Za-454-2 and the wild-type FH9, a low-protein mutant derived from FH9 mutagenized population was also cultivated for analysis.

Follow-up experiments to further evaluate seed protein content at various developmental stages were conducted at both the Breeding Base and the Sihewa Base.

The soil types of both the Breeding Base and the Sihewa Base are Alfisols. Consistent field management practices were implemented across both locations. For each hectare, 750 kg of compound fertilizer (N: P<sub>2</sub>O<sub>5</sub>: K<sub>2</sub>O=15:15:15, Shandong Nongda Fertilizer Technology Co., Ltd., Shandong, China) and 150 kg of medium-element fertilizer

(granular,  $\text{CaO} \geq 9.0\%$ ,  $\text{MgO} \geq 18.0\%$ ) (Sinochem Yantai Crop Nutrition Co., Ltd., Shandong, China) were thoroughly mixed and applied as basal fertilizers.

Peanut was dibbled manually in early May on raised seedbeds 0.9 m wide with double rows, an intra-row spacing of 35 cm, and an inter-plant spacing of 13 cm. The crop was grown without mulching, and soil moisture was initially established through drip irrigation after sowing. Rainfall was sufficient for the remainder of the growing season. In late July, a single side-dressing was performed, with 300 kg of urea (Shandong Yangmei Pingyuan Chemical Co., Ltd., Shandong, China) and 75 kg of potassium sulfate (Binzhou Desheng Chemical Co., Ltd., Shandong, China) applied per hectare.

No fungicides were used for disease management. However, insect pest control measures were implemented during the mid-growing season. Leaf-feeding pests were managed with a 2,000-fold diluted 2.5% beta-cypermethrin emulsion (Shandong Diamond Biological Technology Co., Ltd., Shandong, China) and a 2,000–3,000-fold diluted 1.8% abamectin emulsion (Guangdong Zhongxun Agricultural Science Co., Ltd., Guangdong, China), applied every 7 days for 2–3 cycles. Sucking pests, including thrips, aphids, and

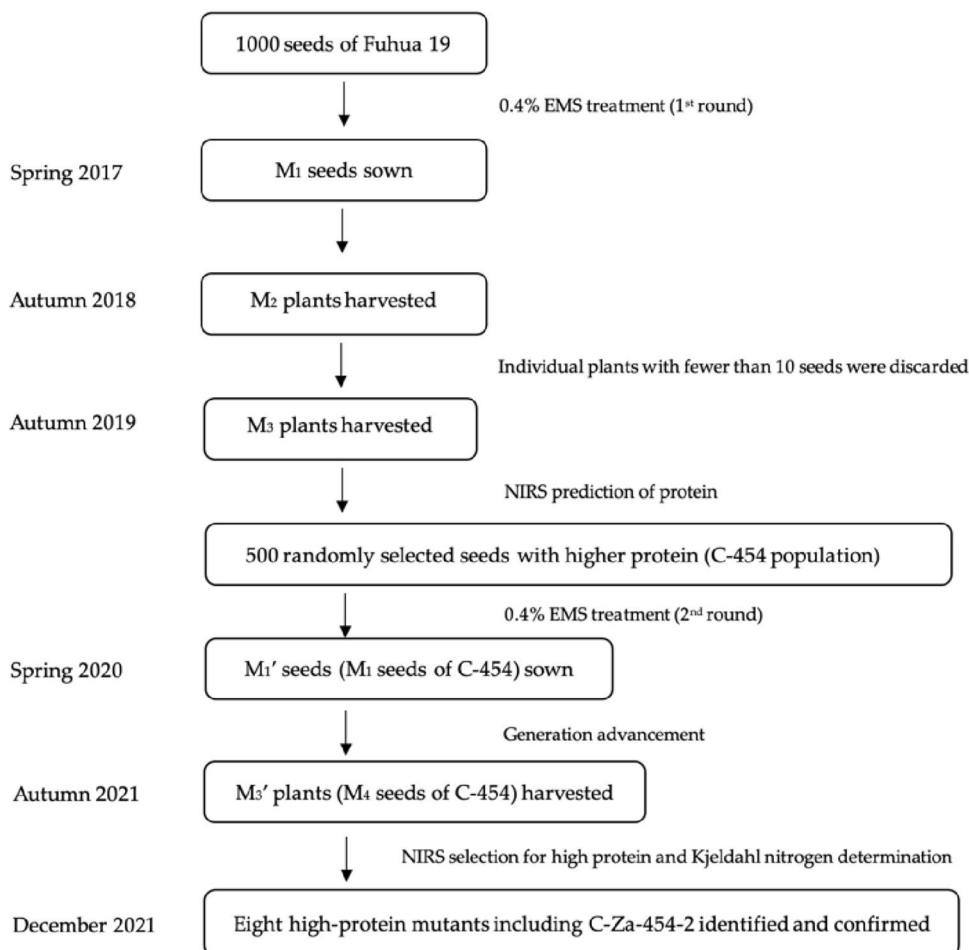
mites, were controlled using 150–225 g per hectare of 20% imidacloprid soluble liquid (Jiangyin No. 2 Pesticide Co., Ltd., Jiangsu, China), a 1,000-fold diluted 20% carbosulfan solution (Shandong Shengbang Lunan Pesticide Co., Ltd., Shandong, China), and 25% thiamethoxam water-dispersible granules (Baoding Shunnong Plant Protection Co., Ltd., Hebei, China), applied every 7 days for 2–3 cycles.

Uplifting and threshing were conducted manually in autumn, after which the pods were air-dried in a drying yard. The dried peanut pods were then hand-shelled.

### Chemical mutagenesis and selection for high protein content

The two-step mutagenesis and selection were carried out during 2017–2021 (Fig. 1). Following a protocol of our research team (Yu et al. 2019), 1000 sound mature seeds of FH19 were treated with 0.4% EMS (Solarbio, Beijing, China) for 6 h to induce mutations.  $M_3$  plants with no less than 10 seeds per plant were selected, while low-yielding plants were discarded. Quality traits of  $M_4$  bulk seeds harvested from the selected  $M_3$  single plants were analyzed using near-infrared (NIR) spectroscopy (NIRS) on a NIR

**Fig. 1** Induction of high-protein peanut mutants with 2 rounds of EMS treatment and NIRS-aided selection



instrument (Matrix I, Bruker Optics, Germany) (Wang and Zhang 2013; Wang et al. 2013) (Fig. 1). A total of 500 randomly seeds with protein content exceeding that of the wild type ( $\geq 25\%$ ) selected from 500 single plants (one seed per plant) were subjected to a second 0.4% EMS treatment (Fig. 1). Selection was made based on productivity, and quality analysis of the second round  $M_4$ ' bulk seeds were conducted using NIR spectroscopy (Wang and Zhang 2013) (Fig. 1).

### Estimation of chlorophyll content

Peanut was sown at the Breeding Base on May 11, 2023. Chlorophyll content was estimated using a Soil Plant Analysis Development (SPAD) chlorophyll meter (SPAD-502PLUS, Konica Minolta, Osaka, Japan) on the third leaf from the top of the main stem during the morning period (9:00–11:00 a.m.) (Nageswara Rao et al. 2001) at six stages (S1: June 26, S2: July 11, S3: July 26, S4: August 10, S5: August 25, S6: September 9, 2023). Each measure was performed in triplicate.

### Sampling of leaves and measurement of biochemical indicators

Peanut was sown on May 11, 2023. Sampling began at the flowering & pegging stage, with peanut leaves (third from the top) collected from the main stem every 15 days across six stages identical to those in chlorophyll estimation (see "Estimation of chlorophyll content" subsection) between 9:00 AM and 11:00 AM. Samples were immediately placed in a liquid nitrogen tank and transported to the laboratory, where they were stored at  $-80^\circ\text{C}$  for subsequent enzyme activity and malondialdehyde (MDA) assay.

Ten indicators associated with nitrogen and carbon metabolism were analyzed. Eight of these were assayed using reagent kits from Nanjing Jiancheng Bioengineering Institute, Nanjing, China, including nitrate reductase (NR), glutamate dehydrogenase (GDH), glutamine synthetase (GS), glutamic-pyruvic transaminase (GPT), glutamic-oxaloacetic transaminase (GOT), sucrose synthase (SS), sucrose phosphate synthase (SPS), and phosphoenolpyruvate carboxylase (PEPCase) activities. The remaining two indicators, glutamate synthase (GOGAT) and ribulose-1,5-bisphosphate carboxylase (RuBPCase) activities, were assessed using reagent kits from Solarbio Science & Technology Co., Ltd.

Four indicators related to peanut leaf senescence were measured using reagent kits provided by Nanjing Jiancheng Bioengineering Institute, Nanjing, China. These included: superoxide dismutase (SOD) activity, peroxidase (POD) activity, catalase (CAT) activity, and MDA content.

Each test was performed in three biological replicates, and the mean values were calculated where necessary.

### Evaluation of protein content in the high-protein mutant C-Za-454-2 and wild-type FH19 across two locations

The high-protein peanut mutant C-Za-454-2 and FH19 were sown on May 15, 2023, at the Breeding Base and the Sihewa Base.

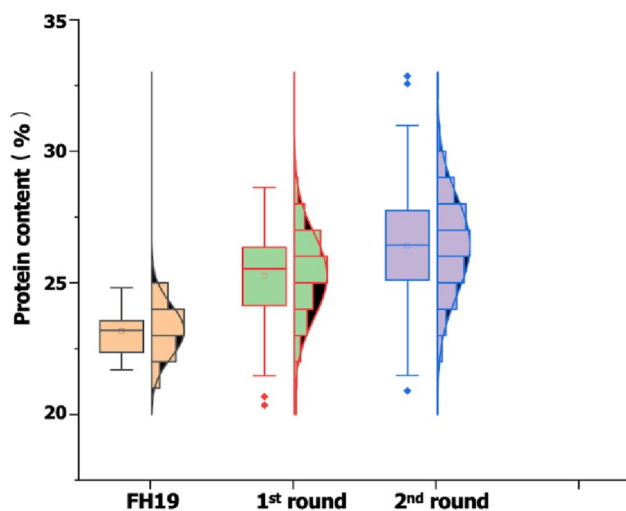
Pegs, two centimeters in length, from both the first and second pairs of lateral branches, were tagged on July 10, 2023. To ensure uniform timing of peg penetration into the soil, U-shaped metal wires were used to facilitate peg entry. Six plants each of the high-protein mutant C-Za-454-2 and the wild-type FH19 were utilized for sampling. Following peg soil penetration, pods were sampled at 10-day intervals on August 4 (P1), 14 (P2) and 24 (P3), and September 4 (P4) and 14 (P5) of the same year.

Protein content of the seed samples was analyzed using the Kjeldahl nitrogen determination method (China standard methods GB 5009.5–2016) (National Health and Family Planning Commission of the People Republic of China, State Administration of Food and Drug Administration 2016).

### Microstructure of peanut cotyledons

Mature peanut seeds were sampled from Breeding Base, cut into  $1\text{ mm}^3$  blocks and immediately immersed in a 2.5% glutaraldehyde solution at  $4^\circ\text{C}$  for 24 h for fixation. The samples were rinsed with phosphate buffer (0.2 mol/L, pH 6.8) for three times and then fixed with a 2% osmium tetroxide solution prepared in phosphate buffer (0.2 mol/L, pH 7.2) for 1 h. After fixation, the samples were rinsed again with phosphate buffer for three times for 15 min each. The samples underwent sequential dehydration using ethanol solutions of increasing concentrations (50%, 70%, 90%, and 100%), with each concentration applied twice for 15 min per step. They were further treated with a 1:1 mixture of anhydrous ethanol and acetone at  $4^\circ\text{C}$ , followed by pure acetone at room temperature for 5 min. The dehydrated samples were infiltrated with SPI812 epoxy resin (Head Biotechnology, Beijing, China) and polymerized in an incubator for 48 h.

Ultrathin sections were prepared using a Leica UC7 ultramicrotome (Leica, Germany) equipped with an Ultra 45° diamond knife (Diatome, USA). The sections were double-stained with 2% uranyl acetate and 2.6% lead citrate solutions at  $25^\circ\text{C}$ . The ultrastructure of the cotyledons was subsequently observed under a Hitachi HT7700 transmission electron microscope (100 kV) (Hitachi, Japan).



**Fig. 2** Range of protein content of the wild-type FH19 and populations from the first and second rounds of mutagenesis after productivity-based selection (bulk seeds from  $M_3$  and  $M_3'$  plants). The boxes represent the interquartile range (IQR), the horizontal line inside the box represents the median, and the whiskers represent the range within 1.5 times the IQR. Outliers are shown as individual points. Error bars are included to indicate standard error of the mean (SEM). The protein content variation data presented for both the primary mutagenized population ( $M_3$ ) (1st round) and secondary mutagenized population ( $M_3'$ ) (2nd round) reflect values prior to NIR selection in their respective generations. However, it should be noted that the development of the  $M_3'$  population was based on NIR-selected individuals from the preceding  $M_3$  generation (500 randomly selected seeds with protein content exceeding that of the wild type FH19, selected from  $M_3$  500 plants)

## Statistical analysis

Analysis of variance (ANOVA), t-test, Tukey's multiple comparison, and correlation analysis were performed using the DPS 14.50 package (Tang and Zhang 2013).

## Results

### Variation in protein content of FH19 and its mutagenized populations post selection

Protein content of wild-type FH19 and its selected mutagenized populations exhibited significant variation (Fig. 2). For the wild type, protein content ranged from 21.70% to 24.82%, with a mean  $\pm$  Standard Deviation (SD) of  $23.18\% \pm 1.00\%$  (Fig. 2). The first round of mutagenesis coupled with selection for plants with more than 10 seeds (787  $M_3$  single plants) resulted in a broader range of 20.35% to 28.62% and a higher mean  $\pm$  SD of  $25.27\% \pm 1.61\%$  (Fig. 2). The second round ( $M_3'$  single plants) further increased both the range, from 20.90% to 32.86%, and the mean  $\pm$  SD to  $26.41\% \pm 1.86\%$  (Fig. 2). These results demonstrated that mutagenesis

**Table 1** Eight high-protein peanut mutants identified from second round of mutagenesis

| High-protein mutant and wild-type FH19 | Protein content (%)* |
|----------------------------------------|----------------------|
| High-protein peanut mutant             |                      |
| C-Za-454-1                             | 30.14                |
| C-Za-454-2                             | 30.32                |
| C-Za-454-3                             | 30.37                |
| C-Za-454-4                             | 30.39                |
| C-Za-454-5                             | 30.9                 |
| C-Za-454-6                             | 30.98                |
| C-Za-454-7                             | 32.57                |
| C-Za-454-8                             | 32.86                |
| FH19 (wild type)                       | 21.70% to 24.82%     |

\*NIRS predicted values

combined with selection significantly expanded the range and elevated the mean protein content, with the second round surpassing the first, and the first exceeding the wild type (Fig. 2). Collectively, these findings highlight the effectiveness of iterative mutagenesis and selection in enhancing protein levels in peanut.

## Identification of high-protein peanut mutants

Through two rounds of mutagenesis followed by selection, eight mutants with protein content exceeding 30% were identified using NIRS (Table 1).

Among these, the mutant C-Za-454-2, when cultivated at the Sihewa Base, displayed a protein content of  $30.23\% \pm 1.21\%$  (mean  $\pm$  SD), which was significantly higher than that of the wild-type FH19 ( $23.07\% \pm 0.67\%$ ) ( $p < 0.01$ ) during the primary evaluation. Additionally, this mutant exhibited a substantially reduced plant height (26.5 cm) compared to FH19 (57.3 cm) and displayed noticeably greener foliage (Fig. 3).

### Protein content of the high-protein mutant C-Za-454-2 and the wild-type FH19 grown at two locations

ANOVA was conducted to examine protein content across five developmental phases, two locations, and two genotypes. The results demonstrated significant differences among genotypes (F-statistic = 1280.464,  $p < 0.01$ ), between locations (F-statistic = 7.2732,  $p = 0.0194$ ), and their interaction (F-statistic = 5.6262,  $p = 0.0353$ ). Temporal variations in protein content were also statistically significant (F-statistic = 5.2372,  $p = 0.0112$ ). Among these factors, genotype emerged as the most influential determinant of protein content.

When averaged across all phases, the wild-type FH19 had a mean protein content of  $23.15\% \pm 0.52\%$  (mean  $\pm$  SD) at the Breeding Base, compared to  $30.35\% \pm 0.96\%$  for

**Fig. 3** Wild-type peanut FH19 (a) and its high-protein mutant C–Za–454–2 (b)



the high-protein mutant C–Za–454–2. At the Sihewa Base, FH19 exhibited a mean protein content of  $23.22\% \pm 0.82\%$ , while the mutant recorded a higher mean of  $31.44\% \pm 0.82\%$ . Notably, the high-protein mutant consistently exhibited protein content exceeding 29% across both locations, outperforming FH19 by more than 6% points on average.

At the maturity phase, the difference between the two genotypes was particularly pronounced. Across both locations, the high-protein mutant maintained an average protein content of 31.32%, whereas FH19 averaged 23.17%. Furthermore, the mutant consistently achieved protein levels above 31% at maturity, while FH19 never exceeded 23.50%. This represents a minimum difference of 7.58% points in favor of the high-protein mutant.

These findings underscore the superiority of the mutant C–Za–454–2 in terms of protein content, with consistent performance across locations and developmental stages.

#### Differences in ultrastructure of cotyledon tissue between the high-protein mutant C–Za–454–2 and the wild-type FH19

Mature seeds of the wild-type peanut FH19 and the high-protein mutant C–Za–454–2 were harvested at the Breeding Base. A portion of the seeds was dried for protein analysis, while another portion of fresh peanut samples was prepared for electron microscopy. The ultrastructure of the cotyledons was illustrated in Fig. 4. The number and area of

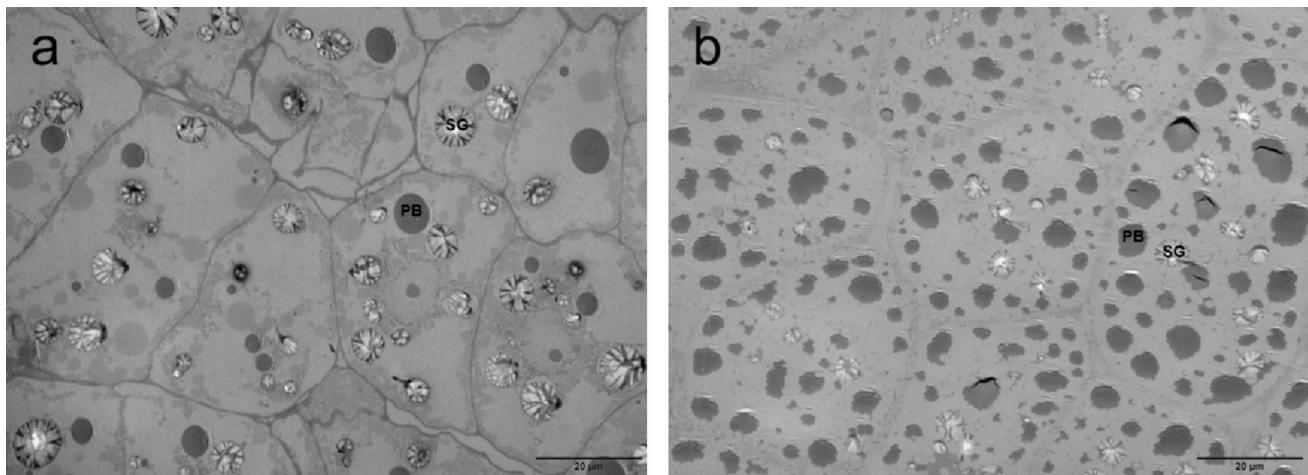
protein bodies and starch grains were quantified. Among the ultrastructural features listed in Table 2, significant differences were observed between the high-protein mutant and the wild-type FH19 in only four parameters: number of protein bodies, total protein body area, number of starch grains, and total starch grain area, at 0.01 level. The high-protein mutant exhibited significantly higher number of protein bodies and total protein body area compared to the wild-type FH19 ( $p < 0.01$ ) (Table 2).

#### Correlation of protein content and the number and area of protein bodies and starch granules

Several statistically significant relationships between protein content and ultrastructural features at 0.01 level were identified through correlation analysis (Table 3).

Protein content (Pro) exhibited a strong positive correlation with number of protein bodies ( $r = 0.90$ ), indicating that higher protein levels were associated with increased number of protein bodies (Table 3). Similarly, total protein body area was strongly positively correlated with protein content ( $r = 0.97$ ), suggesting that greater total area occupied by protein bodies corresponded to higher protein levels (Table 3).

Conversely, a significant negative correlation was observed between number of starch grains and protein content ( $r = -0.91$ ), implying that protein levels increased as number of starch grains decreased (Table 3). Total starch grain area also showed a strong negative correlation with



**Fig. 4** Transmission electron micrographs of cotyledon tissues of FH19 (a) and C-Za-454-2 (b). Scale bar = 20  $\mu\text{m}$ , PB: protein body, SG: starch grain

**Table 2** Number and area of protein bodies and starch grains in seed cotyledonary tissue, expressed as mean  $\pm$  SD

| Anatomical feature                                           | FH19               | C-Za-454-2           |
|--------------------------------------------------------------|--------------------|----------------------|
| Number of protein bodies (No PB)**                           | 24.00 $\pm$ 5.56   | 162.00 $\pm$ 43.18   |
| Minimum protein body area (Min PBA) ( $\mu\text{m}^2$ )      | 1.50 $\pm$ 0.46    | 1.39 $\pm$ 0.51      |
| Maximum protein body area (Max PBA) ( $\mu\text{m}^2$ )      | 40.23 $\pm$ 4.32   | 47.81 $\pm$ 7.13     |
| Average area per protein body (Aver PBA) ( $\mu\text{m}^2$ ) | 15.18 $\pm$ 2.17   | 12.00 $\pm$ 4.28     |
| Total protein body area (Tot PBA) ( $\mu\text{m}^2$ )**      | 356.3 $\pm$ 29.93  | 1826.67 $\pm$ 122.12 |
| Number of starch grains (No SG)**                            | 58.33 $\pm$ 7.76   | 30.33 $\pm$ 3.51     |
| Maximum starch grain area (Max SGA) ( $\mu\text{m}^2$ )      | 1.19 $\pm$ 0.48    | 1.29 $\pm$ 0.91      |
| Maximum starch grain area (Min SGA) ( $\mu\text{m}^2$ )      | 47.73 $\pm$ 6.69   | 33.09 $\pm$ 7.59     |
| Average starch grain area (Aver SGA) ( $\mu\text{m}^2$ )     | 14.87 $\pm$ 3.24   | 14.02 $\pm$ 3.45     |
| Total starch grain area (Tot SGA) ( $\mu\text{m}^2$ )**      | 852.84 $\pm$ 94.69 | 418.15 $\pm$ 59.46   |

Double star (\*\*) denoted significant difference ( $p < 0.01$ )

**Table 3** Correlation of protein content with the number and area of protein bodies and starch grains

|          | No PB   | min PBA | Max PBA | Aver PBA | Tot PBA | No SG   | Min SGA | Max SGA | Aver SGA | Tot SGA |
|----------|---------|---------|---------|----------|---------|---------|---------|---------|----------|---------|
| No PB    | 1.00    |         |         |          |         |         |         |         |          |         |
| min PBA  | 0.00    | 1.00    |         |          |         |         |         |         |          |         |
| Max PBA  | 0.54    | 0.37    | 1.00    |          |         |         |         |         |          |         |
| Aver PBA | -0.47   | -0.74   | 0.72    | 1.00     |         |         |         |         |          |         |
| Tot PBA  | 0.97**  | 0.91**  | -0.10   | 0.67     | 1.00    |         |         |         |          |         |
| No SG    | -0.91** | -0.91** | 0.13    | -0.62    | 0.45    | 1.00    |         |         |          |         |
| Min SGA  | -0.03   | 0.33    | -0.37   | -0.20    | -0.41   | 0.01    | 1.00    |         |          |         |
| Max SGA  | -0.72   | -0.70   | -0.05   | -0.62    | 0.20    | -0.79*  | 0.66    | 1.00    |          |         |
| Aver SGA | -0.10   | -0.11   | -0.01   | -0.16    | -0.02   | -0.19   | -0.12   | -0.19   | 1.00     |         |
| Tot SGA  | -0.93** | -0.91** | 0.07    | -0.57    | 0.47    | -0.96** | 0.65*   | -0.21   | 0.63     | 1.00    |

Star (\*) and double star (\*\*) denoted significance at 0.05 and at 0.01 level, respectively. The full terms for the abbreviations can be found in Table 2

protein content ( $r=-0.93$ ), reinforcing the relationship between reduced starch accumulation and higher protein content (Table 3). Unsurprisingly, the total protein body area showed a significantly strong negative correlation with the total starch grain area or the number of starch grains ( $p<0.01$ ) (Table 3).

### SPAD readings of the high-protein mutant C-Za-454-2 and the wild type FH19 across six stages

As shown in Table 4, the SPAD readings of the high-protein mutant were significantly higher than those of the wild type FH19 ( $p<0.01$  or  $p<0.05$ ) during three stages (S1, S4–S6). However, the differences between the two were not significant during S2 and S3.

Differences in leaf enzyme activities and MDA between high-protein mutant C-Za-454-2 and wild-type FH19 at different growth stages.

As shown in Table 5, among the nitrogen metabolism enzymes, NR (S2–S6), GOGAT (S2–S6), and GS (S1–S3, S5–S6) exhibited significantly higher enzyme activities in the high-protein mutant C-Za-454-2 compared to the wild type FH19 ( $p<0.01$  or  $p<0.05$ ) during five periods. GOT (S2, S6) showed this difference during two periods, while GDH (S6) and GPT (S5) exhibited it during one period each.

In contrast, for the carbon metabolism enzymes, the mutant C-Za-454-2 exhibited significantly lower enzyme activities compared to the wild type FH19 ( $p<0.01$  or  $p<0.05$ ). This was observed for Rubisco (S1, S4, S5) during three periods and for SPS (S2) during one period.

Among the protective enzymes, the mutant displayed significantly lower enzyme activities compared to the wild type FH19 ( $p<0.01$  or  $p<0.05$ ) for CAT (S2, S4, S6) during three periods and for POD (S2, S5) during two periods. However, for SOD, the mutant C-Za-454-2 showed significantly higher enzyme activity than FH19 at the 0.01 level during one period (S1).

Additionally, for MDA levels, the mutant C-Za-454-2 exhibited significantly higher levels compared to the wild type FH19 ( $p<0.05$ ) during two periods (S3 and S4).

### Correlation between protein content of mature peanut seeds and enzyme activities/MDA

The correlation coefficients between the activities of six nitrogen metabolism-related enzymes across six growth stages and the protein content of mature seeds were all positive (Table 6). The enzyme activities of NR and GOGAT were significantly correlated with the protein content of mature seeds across all stages ( $p<0.01$  or  $p<0.05$ ). The correlation coefficients for both enzymes were consistently above 0.69, reaching as high as 0.97 (Table 6). NR exhibited the strongest correlations during the S2–S4 stages, while GOGAT peaked during the S3–S4 stages (Table 6). GS was not significantly correlated during the S4 stage but showed strong positive correlations during other stages, with coefficients ranging from 0.86 (S5) to 0.99 (S6); Notably, the correlations in S1 and S3 were both no less than 0.93. GOT and GDH showed significant correlations (0.05 or 0.01 level) in four stages each. GOT had exceptionally high correlations during the S2 and S6 stages, both exceeding 0.92. GDH demonstrated higher correlations in S6 and S3, both above 0.80 (Table 6).

For carbon metabolism-related enzymes, the activities of SS, SPS, and Rubisco were negatively correlated with the protein content of mature seeds at the 0.05 or 0.01 level (Table 6). Rubisco had a particularly strong negative correlation in the S5 stage, with a coefficient of  $-0.96$ . PEPCase was significantly positively correlated with protein content only in the S4 stage (Table 6).

Among the protective enzymes, POD showed consistent negative correlations with protein content of mature seeds across all stages, with coefficients ranging from  $-0.82$  to  $-0.68$ ; SOD had positive correlations in S1 (0.73), S2 (0.73), and S6 (0.94), but a negative correlation in S5 ( $-0.67$ ); CAT exhibited a negative correlation in S4 ( $-0.88$ ) but positive correlations in S5 (0.71) and S6 (0.84) (Table 6).

MDA was significantly negatively correlated with protein content across the S2–S6 stages, with coefficients ranging from  $-0.85$  to  $-0.68$  at the 0.05 or 0.01 level (Table 6).

**Table 4** SPAD readings between C-Za-454-2 and FH19 at different stages

| Genotype   | SPAD reading |             |             |             |              |              |
|------------|--------------|-------------|-------------|-------------|--------------|--------------|
|            | S1           | S2          | S3          | S4          | S5           | S6           |
| FH19       | 40.12±0.25B  | 41.67±1.23a | 41.99±0.87a | 41.21±0.27b | 36.84±0.32B  | 36.27±0.75B  |
| C-Za-454-2 | 40.71±0.06 A | 42.58±0.53a | 41.57±0.21a | 42.27±0.12a | 39.29±1.07 A | 38.62±0.34 A |

In each column, the values followed by the same lowercase and uppercase letter for FH19 and the high-protein mutant indicated no significant difference at 0.05 level and 0.01 level, respectively

**Table 5** Difference in 14 biochemical indicators in leaves at six stages between high-protein mutant C-Za-454-2 and wild-type FH19

| Biochemical indicator           | Genotype            | S1                   | S2                   | S3                  | S4                   | S5                   | S6                  |
|---------------------------------|---------------------|----------------------|----------------------|---------------------|----------------------|----------------------|---------------------|
| NR ( $\mu\text{mol/h/g}$ )      | FH19                | 0.75 $\pm$ 0.07a     | 0.37 $\pm$ 0.01b     | 0.30 $\pm$ 0.02b    | 0.32 $\pm$ 0.02b     | 0.28 $\pm$ 0.04a     | 0.16 $\pm$ 0.02B    |
| GDH (U/g)                       | High-protein mutant | 0.98 $\pm$ 0.15a     | 0.53 $\pm$ 0.02a     | 0.43 $\pm$ 0.01a    | 0.43 $\pm$ 0.00a     | 0.33 $\pm$ 0.01a     | 0.21 $\pm$ 0.01 A   |
|                                 | FH19                | 0.34 $\pm$ 0.04a     | 0.26 $\pm$ 0.02a     | 0.33 $\pm$ 0.05a    | 0.36 $\pm$ 0.02a     | 0.33 $\pm$ 0.01a     | 0.28 $\pm$ 0.01b    |
| GS ( $\mu\text{mol/h/g}$ )      | High-protein mutant | 0.40 $\pm$ 0.06a     | 0.31 $\pm$ 0.04a     | 0.40 $\pm$ 0.03a    | 0.45 $\pm$ 0.31a     | 0.38 $\pm$ 0.07a     | 0.35 $\pm$ 0.03a    |
|                                 | FH19                | 1.97 $\pm$ 0.25B     | 6.59 $\pm$ 1.09B     | 10.49 $\pm$ 0.25b   | 7.72 $\pm$ 0.34a     | 4.54 $\pm$ 0.84B     | 3.80 $\pm$ 0.23b    |
| GOGAT (U/g)                     | High-protein mutant | 6.62 $\pm$ 1.60 A    | 9.36 $\pm$ 0.29 A    | 13.31 $\pm$ 0.04a   | 8.54 $\pm$ 0.87a     | 6.64 $\pm$ 0.04 A    | 7.21 $\pm$ 0.17a    |
|                                 | FH19                | 293.82 $\pm$ 22.77a  | 348.29 $\pm$ 11.58b  | 590.31 $\pm$ 48.14b | 361.69 $\pm$ 15.31b  | 273.28 $\pm$ 19.44B  | 247.73 $\pm$ 25.16B |
| GOT (U/g)                       | High-protein mutant | 337.58 $\pm$ 10.02a  | 414.38 $\pm$ 4.55a   | 885.02 $\pm$ 39.40a | 575.13 $\pm$ 33.49a  | 399.20 $\pm$ 53.04 A | 314.36 $\pm$ 9.11 A |
|                                 | FH19                | 60.86 $\pm$ 23.26a   | 53.24 $\pm$ 0.72b    | 45.40 $\pm$ 5.18a   | 48.56 $\pm$ 5.91a    | 47.06 $\pm$ 3.12a    | 48.74 $\pm$ 6.56B   |
| GPT (U/g)                       | High-protein mutant | 78.63 $\pm$ 1.24a    | 71.67 $\pm$ 3.57a    | 56.33 $\pm$ 2.84a   | 52.96 $\pm$ 1.09a    | 65.80 $\pm$ 14.26a   | 67.47 $\pm$ 3.51 A  |
|                                 | FH19                | 53.69 $\pm$ 10.24a   | 38.58 $\pm$ 2.77a    | 39.70 $\pm$ 14.45a  | 21.88 $\pm$ 1.77a    | 38.74 $\pm$ 1.76B    | 29.47 $\pm$ 2.34a   |
| SS ( $\mu\text{g/min/g}$ )      | High-protein mutant | 63.22 $\pm$ 7.24a    | 51.36 $\pm$ 14.30a   | 59.90 $\pm$ 21.49a  | 33.06 $\pm$ 7.77a    | 58.60 $\pm$ 8.15 A   | 36.18 $\pm$ 4.11a   |
|                                 | FH19                | 185.42 $\pm$ 12.84a  | 331.25 $\pm$ 18.40a  | 447.92 $\pm$ 61.73a | 435.42 $\pm$ 56.21a  | 310.42 $\pm$ 33.98a  | 239.58 $\pm$ 55.10a |
| SPS ( $\mu\text{g/min/g}$ )     | High-protein mutant | 162.50 $\pm$ 13.50a  | 333.33 $\pm$ 11.79a  | 429.17 $\pm$ 53.11a | 447.92 $\pm$ 84.83a  | 308.33 $\pm$ 48.23a  | 212.50 $\pm$ 27.00a |
|                                 | FH19                | 408.13 $\pm$ 27.39a  | 466.58 $\pm$ 0.41 A  | 449.38 $\pm$ 39.14a | 451.88 $\pm$ 50.34a  | 476.04 $\pm$ 55.25a  | 351.04 $\pm$ 10.97a |
| Rubisco (U/g)                   | High-protein mutant | 358.96 $\pm$ 44.31a  | 410.83 $\pm$ 26.76B  | 420.83 $\pm$ 38.22a | 398.54 $\pm$ 38.87a  | 407.29 $\pm$ 39.71a  | 380.00 $\pm$ 34.37a |
|                                 | FH19                | 93.64 $\pm$ 8.22 A   | 108.93 $\pm$ 10.73a  | 112.76 $\pm$ 8.22a  | 244.62 $\pm$ 20.45 A | 162.96 $\pm$ 2.42a   | 78.36 $\pm$ 12.01a  |
| PEPCase ( $\mu\text{g/min/g}$ ) | High-protein mutant | 58.29 $\pm$ 7.52B    | 91.73 $\pm$ 10.73a   | 91.73 $\pm$ 14.24a  | 170.09 $\pm$ 23.56B  | 112.76 $\pm$ 8.22b   | 94.60 $\pm$ 10.73a  |
|                                 | FH19                | 173.23 $\pm$ 33.37a  | 219.09 $\pm$ 7.66a   | 389.76 $\pm$ 80.66a | 308.56 $\pm$ 22.97a  | 380.64 $\pm$ 67.06a  | 281.49 $\pm$ 33.37a |
| SOD (U/min/g)                   | High-protein mutant | 135.33 $\pm$ 7.66a   | 235.28 $\pm$ 5.88a   | 303.15 $\pm$ 62.66a | 443.89 $\pm$ 100.40a | 384.35 $\pm$ 7.66a   | 303.15 $\pm$ 42.62a |
|                                 | FH19                | 157.64 $\pm$ 22.45B  | 256.56 $\pm$ 14.49a  | 196.72 $\pm$ 32.07a | 207.54 $\pm$ 48.20a  | 146.80 $\pm$ 11.59a  | 63.57 $\pm$ 5.09b   |
| POD (U/min/g)                   | High-protein mutant | 219.87 $\pm$ 12.09 A | 280.17 $\pm$ 27.37a  | 232.08 $\pm$ 3.39a  | 212.41 $\pm$ 17.79a  | 133.45 $\pm$ 7.28a   | 122.64 $\pm$ 6.89a  |
|                                 | FH19                | 482.59 $\pm$ 74.43a  | 653.70 $\pm$ 14.30 A | 546.11 $\pm$ 58.59a | 480.56 $\pm$ 46.94a  | 593.33 $\pm$ 30.94 A | 305.19 $\pm$ 39.85a |
| CAT (U/min/g)                   | High-protein mutant | 367.96 $\pm$ 39.15a  | 608.33 $\pm$ 5.15B   | 424.81 $\pm$ 25.64a | 430.37 $\pm$ 39.00a  | 501.85 $\pm$ 30.41B  | 275.93 $\pm$ 18.68a |
|                                 | FH19                | 7.94 $\pm$ 1.30a     | 13.67 $\pm$ 1.11a    | 10.44 $\pm$ 1.39a   | 9.62 $\pm$ 0.68 A    | 6.78 $\pm$ 1.08a     | 5.72 $\pm$ 0.68B    |
| MDA (nmol/g)                    | High-protein mutant | 8.96 $\pm$ 0.61a     | 8.43 $\pm$ 0.16b     | 9.86 $\pm$ 0.93a    | 7.34 $\pm$ 0.20B     | 9.75 $\pm$ 1.41a     | 8.66 $\pm$ 0.93 A   |
|                                 | FH19                | 26.28 $\pm$ 1.07a    | 26.98 $\pm$ 5.65a    | 39.12 $\pm$ 0.87a   | 82.87 $\pm$ 2.15a    | 72.02 $\pm$ 4.08a    | 82.92 $\pm$ 6.02a   |
| High-protein mutant             |                     | 25.69 $\pm$ 1.70a    | 22.22 $\pm$ 4.67a    | 30.32 $\pm$ 1.18b   | 58.56 $\pm$ 5.96b    | 55.32 $\pm$ 8.59a    | 67.82 $\pm$ 6.27a   |

For the same biochemical indicator at the same growth stage, values followed by the same lowercase letter for FH19 and the high-protein mutant indicated no significant difference at 0.05/0.01 level

**Table 6** Pearson's correlation coefficients between protein content of mature peanut seeds and the 14 biochemical indicators in leaves during six stages

| Biochemical indicator | Correlation coefficient |         |         |         |         |        |
|-----------------------|-------------------------|---------|---------|---------|---------|--------|
|                       | S1                      | S2      | S3      | S4      | S5      | S6     |
| NR                    | 0.79**                  | 0.98**  | 0.97**  | 0.97**  | 0.69*   | 0.91** |
| GDH                   | 0.66*                   | 0.74*   | 0.81**  | 0.28    | 0.59    | 0.83** |
| GS                    | 0.93**                  | 0.89**  | 0.98**  | 0.38    | 0.86**  | 0.99** |
| GOGAT                 | 0.81**                  | 0.74*   | 0.95**  | 0.97**  | 0.80**  | 0.91** |
| GOT                   | 0.58                    | 0.96**  | 0.86**  | 0.63*   | 0.49    | 0.92** |
| GPT                   | 0.62                    | 0.5     | 0.58    | 0.76*   | 0.76**  | 0.33   |
| SS                    | −0.65*                  | −0.42   | −0.31   | 0.07    | 0.2     | 0.01   |
| SPS                   | −0.69*                  | −0.86** | −0.53   | −0.64*  | −0.71*  | −0.23  |
| Rubisco               | −0.93**                 | −0.66*  | −0.72*  | −0.84** | −0.96** | 0.2    |
| PEPCase               | 0.05                    | 0.61    | −0.33   | 0.73*   | 0.51    | 0.51   |
| SOD                   | 0.73*                   | 0.49    | 0.73*   | 0.36    | −0.67*  | 0.94** |
| POD                   | −0.68*                  | −0.78** | −0.82** | −0.61   | 0.08    | 0.04   |
| CAT                   | 0.6                     | −0.56   | −0.52   | −0.88** | 0.71*   | 0.84** |
| MDA                   | −0.61                   | −0.68*  | −0.76*  | −0.85** | −0.85** | −0.73* |

Star (\*) and double star (\*\*) denoted significance at 0.05 and at 0.01 level, respectively. To calculate the correlation coefficient, three groups of peanut materials were used: the wild-type FH19, the high-protein mutant C–Za–454-2, and another FH19 mutagenized line. Each group had three biological replicates at each stage

## Discussion

### Effectiveness of two rounds of mutagenesis in inducing quality variations in peanut

Compared to traits like oil content and oleic acid concentration, protein content in peanut has received relatively little attention in genetic enhancement. This study represents the first successful attempt to create peanut mutants with protein content no less than 30% through two rounds of mutagenesis. The results demonstrate that this technique is a practical and effective strategy for generating high-protein mutants, similar to our previous work on high-oleic peanut breeding (Yu et al. 2019).

Unlike diploid *Arachis* species, the tetraploid cultivated peanut, which originated from the hybridization of two diploid wild species, contains two distinct sub-genomes (A and B). These sub-genomes often harbor duplicate genes that condition the same traits, creating genetic redundancy that can obscure trait variations. Therefore, it is not surprising that two rounds of mutagenesis are more effective in generating diverse quality variations by disrupting redundant or compensatory genetic pathways. Moreover, in the absence of duplicate genes, multiple rounds of mutagenesis can increase the likelihood of inducing mutations in target genes.

### Stability of the high-protein trait

The high-protein trait in the mutant C–Za–454-2 was evaluated across two locations, confirming its stability and heritable nature, as evidenced by the predominant genotype

effect. These findings provide a robust basis for subsequent ultrastructural analysis and enzyme activity assay.

### Potential use of the Dwarf high-protein mutant

In addition to its elevated protein content, the high-protein mutant C–Za–454-2 exhibited reduced plant height, making it a promising parental line for mapping plant height-related genes. Reduced plant height in peanuts offers a significant agronomic advantage by eliminating the need for growth retardants, a common practice to prevent excessive stem elongation and enhance yields. However, improper use of these chemicals can pose food safety risks. The inherent dwarfing observed in this mutant aligns with sustainable agricultural practices, reducing reliance on chemical growth inhibitors.

In 2024, the seed weight of individual plants of the high-protein dwarf mutant and wild-type peanut was measured. The average weight of the mutant was 15.47 g per plant, which, although lower than the wild-type average of 18.57 g per plant, was still in the acceptable range.

### Differences in cotyledon ultrastructure between high-protein mutant and wild-type peanut

In this study, the high-protein mutant exhibited significantly higher protein body numbers and total protein body areas compared to wild-type peanut, while starch grain numbers and total area were reduced. These findings align with those of Zhang et al., who similarly reported an increase in protein body number and area in high-protein peanut (Zhang et al. 2013a) In our previous work, peanut lines with an

average protein content 3% points higher also demonstrated a greater total protein body area, albeit with fewer protein bodies (Li et al. 2024a).

Given that this study utilized the mutant with a protein content exceeding 30% and a genetic background closely resembling the wild type, the conclusions regarding the relationship between protein content and protein body traits are likely more reliable. Notably, the correlation coefficient between protein content and total protein body area (0.97) was higher than that with protein body number (0.90), indicating that total protein body area may play a more substantial role in determining protein content.

### Leaf biochemical indicators and seed protein content

Zhang et al. (2013b) investigated the differences in leaf enzyme activities between the high-protein peanut line KB008 (29.7% protein) and two lower-protein peanut cultivars, Hua 17 (26.1%) and Nongda 818 (26.5%). Their results showed that NR activity was higher in the high-protein line during early stages, with smaller differences later. GS activity remained consistently highest in KB008 across all stages, with the most significant difference occurring during flowering. Similarly, KB008 exhibited higher GDH activity throughout all developmental stages. Differences in GOGAT and GPT activities were most pronounced in the later stages (Zhang et al. 2013b). These findings partially align with our study but also have notable differences. For example, in our research, for NR and GOGAT activity, C-Za-454-2 and FH19 did not differ significantly in the first stage; however, in the subsequent five stages, the high-protein mutant consistently exhibited higher activity than the wild type.

Li et al. (2024b) found that nitrogen application in the podding zone enhanced peanut protein content. Further investigation revealed that nitrogen application upregulated the expression of nitrogen metabolism-related genes in pods, including NR, nitrite reductase (*NIR*), and GS. This was accompanied by increased activities of NR, GS, and GOGAT enzymes in the pods, consistent with the elevated gene expression levels (Li et al. 2024b). Although the organ in their study (pod, for enzyme activity and gene expression analysis) differed from that in our research (leaf, for enzyme activity determination), both studies confirmed that there was a probable link between increased nitrogen metabolism enzyme activity and elevated protein content.

In this study, the findings suggest that the high-protein peanut mutant exhibits enhanced activities of enzymes associated with nitrogen metabolism, which directly correlate with increased protein content in mature seeds. The elevated activities of NR, GS, and GOGAT in the high-protein mutant

indicate a robust nitrogen assimilation pathway, essential for synthesizing amino acids and proteins necessary for seed development. Conversely, the negative correlations observed between certain carbon metabolism enzymes (like SPS and Rubisco) and protein content suggest a complex interplay where excessive carbohydrate synthesis might limit available resources for protein production during critical growth phases.

From another perspective, the strong correlation between several nitrogen metabolism-related enzyme activities and protein content at maturity in this study may reflect the physiological and biochemical mechanisms underlying the major gene controlled high-protein trait of the high-protein mutant.

### Future research needs

Our findings reveal initial correlations; however, larger-scale studies with diverse genetic materials will be critical to establish causative metabolic relationships with seed protein accumulation. While the data from the present study demonstrate strengthened nitrogen assimilation in the high-protein mutant, complementary measurements of nodule activity (e.g., acetylene reduction assays) would be needed to determine whether nitrogen fixation efficiency is concurrently enhanced in this mutant.

Mondal and Badigannavar (2016) previously reported a significant increase in the content of coarachin, a subunit component of soluble proteins, in a high-protein peanut mutant. Earlier studies have also documented the relationship between the amino acid composition and subunit components of peanut protein and its functional properties (Wang 2018). Elucidating the differences in amino acid composition and storage protein subunit profiles between high-protein mutant and wild-type peanuts is of considerable scientific interest. These variations not only directly affect the nutritional quality of peanut protein but also have the potential to influence its functional properties. Of course, it is meaningful to directly explore whether the mutant has undergone beneficial changes in protein functional properties. Reportedly, peanut protein is inferior to soybean protein in terms of gelation properties (Wang 2018), and peanut protein has relatively low solubility (Cui et al. 2023). The elevated levels of essential amino acids, coupled with favorable modifications in storage protein subunit composition and/or functional properties, may significantly enhance the utility of peanut protein in food and feed applications. The high-protein mutant obtained in this study, along with the wild-type peanut, provide valuable parental materials for the mapping of high-protein genes in peanut. The findings from this study hold significant implications for advancing the genetic improvement of peanut varieties for edible uses.

## Conclusions

This study utilized an innovative two-step mutagenesis approach combined with single-plant seed set selection and NIRS screening to successfully develop high-protein peanut mutants with protein content exceeding 30%, derived from the low-protein cultivar FH19 (<25% protein at seed maturity). The mutant demonstrated stable protein content across two locations and exhibited distinct ultrastructure features. Specifically, the high-protein mutant showed significantly greater number and total area of protein bodies in seed cotyledon tissues compared to the wild-type FH19.

Furthermore, the mutant displayed elevated chlorophyll content in leaves at seed maturity stage, along with increased nitrogen metabolism enzyme activities across most growth stages. The activity of nitrogen metabolism-related enzymes in leaves at six different stages exhibited a positive correlation with protein content, while the activity of carbon metabolism-related enzymes generally showed a negative correlation with protein content. The high-protein trait in the mutant is underpinned by unique ultrastructural, physiological, and biochemical characteristics, providing a robust foundation for further investigation into the genetics and underlying genes responsible for high protein content in peanut.

**Acknowledgements** This work was research by the National Natural Science Foundation of China (Grant No. 32202326); Natural Science Foundation of Shandong Province (Grant No. ZR2024MC186), Liaoning Province Germplasm Innovation and Grain Storage Technology Special Program (Grant No. 2023JH110200002), Qingdao Natural Science Foundation (Grant No. 24-4-4-zrjj-38-jch), the Agricultural Scientific and Technological Innovation Project of Shandong Academy of Agricultural Sciences (Grant No. CXGC2024F20), and China Agricultural Research System (Grant No. CARS-13).

**Author contributions** S.T.Y. conceived the study, analyzing the data, obtaining project funding, drafting the initial version of the paper, and revising the paper. X.Y., Y.C.Y. and Y.Z., managed resource selection and development, and drafting the initial version of the paper. J.C.D. and L.L.G. executed SPAD reading measurement, sampling and enzyme assays, and drafting the initial version of the paper. H.J.S. and G.D.Y., conducted NIRS analysis and nitrogen determination, and drafting the initial version of the paper. C.J.J. and B.B.L., in charge of microstructure observations, and drafting the initial version of the paper. C.T.W., responsible for conceiving the study, experimental methods, resource provision, obtaining project funding, supervising the implementation of the research, and revising the paper. J.Y., coordinated data analysis, obtaining project funding, and drafting the initial version of the paper. All authors have agreed to the final version of the manuscript and are willing to take responsibility for the accuracy and authenticity of the entire research work to ensure that any issues related to the accuracy or integrity of any part of the manuscript are appropriately investigated and resolved. All authors have read and agreed to the current version of the manuscript.

**Data availability** No datasets were generated or analysed during the current study.

## Declarations

**Competing interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Chen H, Liu N, Huang L, Huai D et al (2024) Identification of a major QTL for seed protein content in cultivated peanut (*Arachis hypogaea* L.) using QTL-seq. *Plants* 13:2368. <https://doi.org/10.3390/plants13172368>
- Cui S, McClements DJ, Xu X (2023) Peanut proteins: extraction, modifications, and applications: a comprehensive review. *Grain Oil Sci Technol* 6:135–147. <https://doi.org/10.1016/j.gaost.2023.07.001>
- Huang B, Liu H, Fang Y et al (2023) Identification of oil content QTL on Arahyl2 and Arahyl6 and development of KASP markers in cultivated peanut (*Arachis Hypogaea* L.). *J Integr Agric* <https://doi.org/10.1016/j.jia.2023.11.010>
- Janila P, Pandey MK, Shasidhar Y et al (2016) Molecular breeding for introgression of fatty acid desaturase mutant alleles (*ahFAD2A* and *ahFAD2B*) enhances oil quality in high and low oil containing peanut genotypes. *Plant Sci* 242:203–213. <https://doi.org/10.1016/j.plantsci.2015.08.013>
- Kassie FC, Nguempjop JR, Ngalle HB et al (2023) An overview of mapping quantitative trait loci in peanut (*Arachis hypogaea* L.). *Genes* 14:1176. <https://doi.org/10.3390/genes14061176>
- Li L, Lin S, Liao X et al (1994) Study on the quality of seed and hybridization in wild species of genus *Arachis*. *China Oil Crops* 16(2):22–26
- Li J, Jiang CJ, Sun HJ (2024a) Characterization of two peanut genotypes for biochemical quality attributes and cotyledon microstructure features. *Genet Resour Crop Evol* 71:145–155. <https://doi.org/10.1007/s10722-023-01612-w>
- Li G, Guo X, Sun W et al (2024b) Nitrogen application in pod zone improves yield and quality of two peanut cultivars by modulating nitrogen accumulation and metabolism. *BMC Plant Biol* 24:48. <https://doi.org/10.1186/s12870-024-04725-1>
- Liang JN, Zhang J, Tan LZ et al (2022) Transcriptome analysis of peanut protein mutants at different seed development stages. *Chin J Oil Crop Sci* 44:1018–1029. <https://doi.org/10.19802/j.issn.1007-9084.2021257>
- Lusas EW, Erickson DR, Nip W-K (1989) Food uses of whole oil and protein seeds. The American Oil Chemists Society, Champaign, IL, USA
- Mallikarjuna N, Varshney RK (2014) Genetics, genomics and breeding of peanuts. Taylor & Francis Group, LLC, Boca Raton
- Mondal S, Badigannavar AM (2016) Development of high protein groundnut (*Arachis Hypogaea* L.) mutant through induced

- mutagenesis. *J Oilseeds Res* 33:98–101. <https://doi.org/10.56739/jor.v33i2.138886>
- Nageswara Rao RC, Talwar HS, Wright GC (2001) Rapid assessment of specific leaf area and leaf nitrogen in peanut (*Arachis hypogaea* L.) using a chlorophyll meter. *J Agron Crop Sci* 186:175–182. <https://doi.org/10.1046/j.1439-037X.2001.00472.x>
- National Health and Family Planning Commission of the People Republic of China, State Administration of Food and Drug Administration (2016) National standard for food safety—determination of protein in foods. Beijing, China
- Nwokolo E, Smartt J (1996) Food and feed from legumes and oilseeds. Springer US, Boston, MA, USA
- Sakha MA, Jefwa J, Gweyi-Onyango JP (2022) Critical analysis and evaluation of groundnut value chain for revamping its production for global food security. In: Kumar A, Kumar P, Singh SS (eds) Agriculture, Livestock Production and Aquaculture. Springer, Cham, pp 43–58
- Singh B, Singh U (1991) Peanut as a source of protein for human foods. *Plant Foods Hum Nutr* 41:165–177
- Smartt J (1994) The groundnut crop: a scientific basis for improvement. Springer Netherlands, Dordrecht
- Stalker HT (2017) Utilizing wild species for peanut improvement. *Crop Sci* 57:1102–1120. <https://doi.org/10.2135/cropsci2016.09.0824>
- Stalker HT, Wilson RF (2016) Peanuts: genetics, processing, and utilization. Academic Press, Amsterdam, Boston
- Tang Q-Y, Zhang C-X (2013) Data processing system (DPS) software with experimental design, statistical analysis and data mining developed for use in entomological research. *Insect Sci* 20:254–260
- Vargheese RL, Saravanan S, Hepziba SJ et al (2024) Marker-assisted introgression to improve the oleic acid content in the TMV 7 groundnut (*Arachis hypogaea* L.) variety suitable for the oil industry. *BMC Plant Biol* 24:733. <https://doi.org/10.1186/s12870-024-05387-9>
- Varshney RK, Pandey MK, Puppala N (2017) The peanut genome. Springer International Publishing, Cham
- Wang Q (2018) Peanut processing characteristics and quality evaluation. Springer Singapore, Singapore
- Wang CT, Zhang JC (2013) Peanut genetic improvement. Shanghai Sci & Tech Press, Shanghai, China
- Wang CT, Wang XZ, Tang YY (2013) Predicting main fatty acids, oil and protein content in intact single seeds of groundnut by near infrared reflectance spectroscopy. *Adv Mater Res* 860:490–496. <https://doi.org/10.4028/www.scientific.net/AMR.860-863.490>
- Whitworth RJ (2021) *Arachis hypogaea*: cultivation, production and nutritional value. Nova Science Pub Inc, Hauppauge, New York
- Woodroof JG (1983) Peanuts: production, processing, products, 3rd edn. AVI Pub, Westport, Conn
- Yang Y, Li Y, Cheng Z et al (2023) Genetic analysis and exploration of major effect QTLs underlying oil content in peanut. *Theor Appl Genet* 136:1–14. <https://doi.org/10.1007/s00122-023-04328-8>
- Yu ST, Yu HB, Yu GQ et al (2015) Isolation of differentially expressed genes from developing seeds of a high-protein peanut mutant and its wild type using Genefishing™ technology. *Lect Notes Electr Eng* 332:37–46
- Yu ST, Wang CT, Wang H et al (2019) Two-step chemical mutagen treatment to convert Fuhua 12, a normal oleic Spanish peanut cultivar, into its high oleic version. *IOP Conference Series: Earth and Environmental Science* 346:012039. <https://doi.org/10.1088/1755-1315/346/1/012039>
- Zhang JL, Li X-D, Yang CT (2013a) Relationship of quality formation and ultrastructure of cotyledon cells in two quality types of peanut. *Chin J Plant Ecol* 37:768. <https://doi.org/10.3724/SP.J.1258.2013.00080>
- Zhang J-L, Gao F, Lin Y-J (2013b) Differences in seed kernel quality and related enzyme activities of different quality type peanut cultivars. *Chin J Appl Ecol* 24:81–487
- Zhao T, Ying P, Zhang Y, Chen H (2023) Research advances in the high-value utilization of peanut meal resources and its hydrolysates: a review. *Molecules* 28:6862

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.