Metabolic profiling and spatial metabolite distribution in wild soybean (*G. soja*) and cultivated soybean (*G. max*) seeds

Xin Yin [a,1], Zhentao Ren [a,1], Ruizong Jia [b], Xiaodong Wang [c], Qi Yu [a], Li Zhang [a], Laipan Liu [a], Wenjing Shen [a], Zhixiang Fang [a], Jingang Liang [d,*], Biao Liu [a,*]

[a] *Nanjing Institute of Environmental Sciences, Ministry of Ecology and Environment, Nanjing 210042, China*
[b] *Sanya Research Institution/Hainan Key Laboratory for Biosafety Monitoring and Molecular Breeding in Off-Season Reproduction Regions, Chinese Academy of Tropical Agriculture Sciences, Sanya 572011, China*
[c] *Key Laboratory of Mass Spectrometry Imaging and Metabolomics (Minzu University of China), State Ethnic Affairs Commission, Beijing 100081, China*
[d] *Development Center of Science and Technology, Ministry of Agriculture and Rural Affairs, Beijing 100176, China*

ARTICLE INFO

ABSTRACT

Wild soybeans retain many substances significantly reduced or lost in cultivars during domestication. This study utilized LC-MS to analyze metabolites in the seed coats and embryos of wild and cultivated soybeans. 866 and 815 metabolites were identified in the seed extracts of both soybean types, with 35 and 10 significantly differing metabolites in the seed coat and embryos, respectively. The upregulated metabolites in wild soybeans are linked to plant defense, stress responses, and nitrogen cycling. MALDI-MSI results further elucidated the distribution of these differential metabolites in the cotyledons, hypocotyls, and radicles. In addition to their role in physiological processes like growth and response to environmental stimuli, the prevalent terpenoids, lipids, and flavonoids present in wild soybeans exhibit beneficial bioactivities, including anti-inflammatory, antibacterial, anticancer, and cardiovascular disease prevention properties. These findings underscore the potential of wild soybeans as a valuable resource for enhancing the nutritional and ecological adaptability of cultivated soybeans.

1. Introduction

Wild soybean (*Glycine soja* Siebold & Zucc., GS), the progenitor of cultivated soybean (*Glycine* max (L.) Merr., GM), is an annual self-pollinating herb and a critically important strategic biological resource. China is the center of biodiversity for wild soybeans (Dong, Zhuang, Zhao, Sun, & He, 2001), and has classified them as a nationally protected species over the past two decades. The natural distribution of wild soybeans in China highly overlaps with cultivated soybean planting areas, sharing similar growth periods, and exhibiting cross-fertility (Nakayama & Yamaguchi, 2002). Cultivated soybeans are domesticated from wild soybeans through long-term artificial selection (Carter, Hymowitz, & Nelson, 2004; Kim, Van, Kang, Kim, & Lee, 2012). Wild soybeans are characterized by black seed coats and possess smaller seeds while demonstrating superior environmental adaptability compared to cultivated varieties(Y.-h. Li et al., 2014). Wild soybeans typically have higher levels of protein, linolenic acid, stachyose, and raffinose and lower levels of oil and oleic acid, which have positive implications for

human health (Y. Chen & Nelson, 2004; La et al., 2019; Leamy, Zhang, Li, Chen, & Song, 2017).

Prolonged domestication and improvement processes have led to a significant loss of genetic material in wild soybeans, particularly genes related to environmental adaptability (Kofsky, Zhang, & Song, 2018). Wild soybeans exhibit stronger resistance to various biotic and abiotic stresses, such as drought, pests, and salt stress than domesticated varieties (Cai, Jia, Sun, & Sun, 2022; Liu, Li, Liu, & Shi, 2020). These adaptive capabilities are also reflected in the metabolic differences. For instance, a comprehensive metabolic analysis of different varieties of cultivated black and wild soybeans identified 48 metabolites with significant differences in their quantities, including higher levels of flavonoids and phenylpropanoids in wild soybeans (Hyeon, Xu, Kim, & Choi, 2020). Meanwhile, 98 differential metabolites (isoflavones, free amino acids, and fatty acids) were identified from four wild and ten cultivated soybean genotypes using untargeted metabolomics (Tareq, Kotha, Natarajan, Sun, & Luthria, 2023). Additionally, numerous studies have focused on metabolic changes in wild soybeans under biotic and abiotic

---

* Corresponding authors.
   E-mail addresses: liangjingang@agri.gov.cn (J. Liang), liubiao@nies.org (B. Liu).
   [1] X.Y. and Z.R. contributed equally to this work.

stresses (Bao, Mu, & Wang, 2023; J. Chen et al., 2022).

Despite extensive research on soybean domestication and genetic engineering, less attention has been given to the metabolic distinctions between wild and cultivated soybeans, especially in the context of their seed coats and embryos.The seed coat and embryo are crucial components of plant growth and development and play different roles in seed dormancy and developmental processes (Ben-Tov et al., 2015; Lafon-Placette & Köhler, 2014). The seed coat protects the embryo from mechanical damage and pathogen infection, influencing seed viability and the ability to withstand biotic and abiotic stresses, and its permeability is a crucial factor in seed germination (Radchuk & Borisjuk, 2014; Zhou et al., 2022). In addition to storing nutrients for seed development, the embryo produces plant hormones, such as abscisic acid (ABA), that maintain seed dormancy (Wang et al., 2018; H. Zhang et al., 2023). Therefore, a comprehensive understanding of the metabolic components of different parts of wild soybean seeds can reveal the distinct roles of the seed coat and embryo, and uncover unique genetic traits beneficial for soybean variety improvement.

Liquid chromatography-mass coupled with mass spectrometry (LC-MS) is a commonly used technique for metabolomic analysis that is widely applied in various fields (Anagnostopoulos, Stasinopoulou, Kanatas, & Travlos, 2020; Gong et al., 2020; C.-R. Li et al., 2022; Llorach et al., 2019). However, it does not display the spatial distribution of metabolites, whereas matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) reveals the molecular distribution within tissue sections (Fu et al., 2023; Wu et al., 2022). However, previous research has primarily addressed whole soybean seeds, with few reports exploring the metabolic differences in specific parts of soybean seed, such as seed coats and embryos.

This study employed LC-MS to compare the metabolites in the seed coats and embryos of the three wild soybean varieties and three cultivated soybean varieties. Additionally, MALDI-MSI technology was used to display the spatial distribution differences of metabolites in the soybean embryo, including the cotyledons, hypocotyls, and radicles. The objectives of this study were to 1) compare the metabolic profiles of seed coats amd embryos between wild and cultivated soybeans, 2) identify and quantify differential metabolites that contribute to the observed phenotypic and functional disparities between wild and cultivated soybeans, and 3) reveal the spatial distribution of key differential metabolites within seeds, providing insights into their biological roles in situ. These objectives aim to deepen our understanding of how domestication has altered the metabolic landscape of soybeans and to identify specific metabolites that could be leveraged to enhance the nutritional and environmental resilience of cultivated varieties.

2. Materials and methods

2.1. Materials

Three wild soybean varieties (GS-YN, GS-JS, and GS-JL) and three cultivars (GM-ZLD, GM-SD21, and GM-HX10) were collected from three ecological regions in China: North, Central, and South. The morphology and weight of the soybean seeds are shown in supplementary Fig. S1 and Table S1. All the seeds were uniformly bred in Nanjing, Jiangsu Province, China. Three plants were selected per variety for sampling from the resulting progeny. All seeds were refrigerated at 4 °C until further use.

2.2. Chemicals and reagents

All chemicals and solvents used in this study were of high-performance liquid chromatography (HPLC) or analytical grade. Methanol, acetone, sodium hydroxide, hydrochloric acid, formic acid, ethyl acetate, diethyl ether, ethanol, sodium carbonate, sodium nitrite, and aluminum chloride were purchased from Merck (Darmstadt, Germany). The matrix for MALDI-MS, extrapure 2-mercaptobenzothiazole (2-MBT), was procured from Sigma-Aldrich (St. Louis, MO, USA). Ultrapure water (type 1) used throughout the experiments was obtained via a Milli-Q system (Millipore Corporation, Bedford, MA, USA). This stringent adherence to high-purity standards ensured the reliability and reproducibility of the analytical results.

2.3. Metabolite extraction

For metabolite extraction, 200 mg samples of seed coats and embryos from each soybean variety were prepared. The samples were converted into homogenous fine powders by ball milling. Each powdered tissue specimen was transferred to a sterile 2-mL Eppendorf tube. The samples were suspended in 500 μL methanol:water mixture (80:20, v/v), and four 5 mm stainless steel balls were added, followed by homogenization using a Precellys® 24 tissue homogenizer (Bertin Technologies, France) for three intervals of 60 s each.

An additional 500 μL of the methanol/water (80:20, v/v) solution was added to each tube followed by a further 30-s bout of homogenization to ensure thorough mixing and breakdown of cellular matrices. The homogenate was centrifuged at 10000g for 20 min at 4 °C using an Eppendorf 5430 R centrifuge (Hamburg, Germany) to effectively separate the phases. The resultant supernatant containing the organic phase was meticulously aspirated and transferred into a 2-mL Eppendorf tube. The organic extracts were subsequently dehydrated at 4 °C using a Savant SPD111 SpeedVac concentrator (Thermo Scientific, Waltham, MA, USA) to concentrate the metabolites. Following lyophilization, the residual metabolite pellet was resuspended in 500 μL of 50% acetonitrile (ACN) and ultrasonicated for 10 min to facilitate dissolution. The final step entailed filtering the metabolite solution through a 0.22-μm pore-size filter membrane to ensure that the samples were devoid of particulate matter prior to analysis by LC-Tandem MS (LC-MS/MS).

2.4. LC-MS/MS

Metabolomic separation and detection were conducted using a state-of-the-art Waters ACQUITY Ultra Performance Liquid Chromatography (UPLC) system seamlessly interfaced with an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with an electrospray ionization (ESI) source. The chromatographic run was facilitated by a binary gradient elution profile utilizing 0.01% formic acid in water as Solvent A in a 1:1 mixture of ACN and isopropanol as Solvent B. The gradient commenced with a linear increase from 5% to 45% Solvent B over 5 min, followed by an increase to 100% Solvent B over the next 20 min and holding at 100% Solvent B for an additional 10 min. Metabolite separation was achieved using a Waters ACQUITY UPLC BEH C18 column (1.7 μm particle size, 2.1 mm internal diameter × 50 mm length, Beverly, MA). Elution was performed at a controlled flow rate of 0.35 mL/min and column temperature of 45 °C to secure optimal resolution before subsequent injections. Fourier-transform mass spectrometry (FTMS) data acquisition was configured to operate at a mass resolution of 15,000 Full Width at Half Maximum (FWHM) at m/z 400 and scanning a mass range between m/z 30 and 1500 for broad-spectrum metabolite coverage. The MS/MS experiments employed collision-induced dissociation (CID) with a normalized collision energy of 35% to yield information-rich fragmentation patterns. Precision control over ion population was managed by an automatic gain control target of 2×10^5 ion counts, combined with the maximum injection time per ion set at 100 ms in the ion trap to ensure data uniformity. Dynamic exclusion was implemented during MS/MS to optimize the detection of less abundant ions, employing a 15-s exclusion window post-ion selection. Metabolite identification was meticulously performed by cross-referencing the acquired MS/MS spectral profiles against established repositories such as METLIN (https://metlin.scripps.edu/), the Human Metabolome Database (HMDB) (https://hmdb.ca/), and MassBank (https://massbank.eu/MassBank/), supplemented by expert manual spectral interpretation to enhance identification accuracy.

2.5. Tissue sectioning

Uniform soybean seeds of a consistent size were carefully selected for preparation. The chosen seeds were submerged in ice-cold deionized water maintained at 0 °C for 8 h to facilitate easy slicing. The seeds were rapidly transferred to a Leica CM1860 cryostat (Leica Microsystems Inc., Wetzlar, Germany) and cryo-sectioned at −20 °C. The sectioning protocol was carefully aligned horizontally and parallel to the orientation of the soybean seed cotyledons (Fig. S2). We collected serial sections with a thickness of 12 μm near the locus representative of the seeds' maximal cross-sectional girth. These sections were promptly mounted onto the conductive surfaces of indium tin oxide (ITO)-coated glass slides (Bruker Daltonics, Billerica, MA, USA) designed for optimal microscopic analysis. The mounted sections were transitioned to a fume hood to undergo ambient air drying before matrix coating.

2.6. Matrix coating

A 2-MBT matrix solution was prepared at 10 mg/mL in an optimized solvent system of acetonitrile/water (ACN/H_2O, 80:20 *v*/v) containing 0.3% trifluoroacetic acid (TFA) to aid analyte co-crystallization and ionization. Following air-drying of the tissue sections, we employed a precision HIT-MatrixPrep (Huayi Innovation and Biotechnology Co., Ltd., Beijing, China) for matrix deposition. The 2-MBT matrix solution was spray-coated onto an ITO-coated glass slide at a density of ~0.18 mg/cm^2 over 8 min. The flow rate of the matrix solution was 200 μL/min, and the assisted nitrogen gas pressure was maintained at 0.65 MPa to ensure a uniform and comprehensive coating of the tissue sections. High-resolution optical images were acquired using an Epson Perfection V550 Photo Scanner, which provided detailed visual records of sample preparation for subsequent analytical referencing to document and preserve the morphology of the matrix-coated sections.

2.7. Matrix-assisted laser desorption/ionization mass spectrometry

Our analysis employed a Bruker Autoflex Speed MALDI time-of-flight (TOF)/TOF mass spectrometer (Bruker Daltonics, Billerica, MA, USA) for in situ detection and spatial mapping of flavonoids within soybean seed tissue sections. The MALDI source was integrated with a state-of-the-art 2000 Hz Smartbeam Nd:YAG solid-state UV laser operating at a wavelength of 355 nm (Azura Laser AG, Berlin, Germany). For comprehensive MALDI-MS profiling, we collected mass spectra within the *m/z* range of 200–1000 in the positive-ion mode. The refined data acquisition protocol included the accumulation of 100 scanned repetitions, with each scan consisting of 500 laser shots. The laser power was meticulously optimized to 50% of its maximum capacity while factoring in a global attenuator offset of 20%. This calibrated setup delivered an energy output of 1.3 mJ per laser pulse, strictly adhering to the manufacturer's technical specifications. During MALDI-MS imaging (MALDI-MSI), high-quality images were captured with a spatial resolution of 100 μm using 500 laser shots per pixel to elucidate the distribution of endogenous flavonoids. We selected the matrix ion of 2-MBT ([M + H]$^+$, *m/z* 167.99) in conjunction with the well-characterized bradykinin 1–7 peptide standard ([M + H]$^+$, *m/z* 757.40) for precise mass calibration.

2.8. Data analysis

For initial processing, Bruker FlexAnalysis 3.4 software was used for tasks such as mass spectral visualization, peak alignment, batch internal calibration, and discerning monoisotopic peaks displaying a signal-to-noise ratio (S/N) of 3. The latest version of Bruker FlexImaging 4.1 software was employed to construct precise ion maps representing the spatial distribution of endogenous flavonoids that allowed for an exceedingly stringent mass filter width of 10 ppm. The acquired profiling data were uploaded to MetaboAnalyst 4.0, an online platform (https://www.metaboanalyst.ca/), where they underwent rigorous

analysis in accordance with the indicated guidelines. Using this advanced platform, we performed partial least squares discriminant analysis (PLS-DA) to effectively visualize and interpret the amassed MALDI-MS multivariate datasets (Xia, Sinelnikov, Han, & Wishart, 2015).

3. Results and discussion

3.1. Metabolic identification and profiling in wild and cultivated soybean seeds via LC-MS

Metabolic profiling of seed coats and embryos identified a total of 886 metabolites in wild and cultivated varieties, with 722 in the seed coats and 489 in the embryos (Fig. 1A). Specifically, 866 metabolites were detected in wild soybeans, with 703 in the seed coats and 471 in the embryos, whereas 815 metabolites were identified in cultivated soybeans, with 643 in the seed coats and 467 in the embryos. There were 56 metabolites exclusive to wild soybean seed coats and 12 to cultivated soybean seed coats. Additionally, 10 metabolites were exclusive to wild soybean embryos and 7 to cultivated soybean embryos. Soybean seeds shared 795 metabolites, demonstrating a similarity of 89.7%. Furthermore, 280 metabolites were common to the seed coats and embryos of the two soybean types. There were significant differences in the quantities of metabolites between the wild and cultivated soybean seed coats, whereas the quantities in the embryos were nearly identical. The primary metabolites in seed coats and embryos were terpenoids, lipids, and flavonoids (Fig. 1C, D), although their proportions varied. The top three metabolite types in cultivated soybean embryos were terpenoids (16.6%), flavonoids (16.2%), and lipids (14.2%), whereas those in wild soybean embryos were flavonoids (18.2%), terpenoids (15.7%), and lipids (13.4%). Generally, most metabolites in wild soybean seeds were upregulated compared with those in cultivated soybeans. Terpenoids, flavonoids, and lipids were upregulated by 62.1%, 89.1%, and 58.6% in seed coats and 70%, 91.8%, and 56.5% in embryos, respectively, in wild soybean varieties (Fig. S3).

Terpenoids are vital for plant growth and development and are involved in photosynthesis, respiration, growth regulation, and protection against high light stress (Langenheim, 1994; Tholl, 2015). Plants contain abundant flavonoids. Beyond their role in pigment deposition, they regulate cell growth and differentiation, exhibit antioxidant activity, and play a role in pathogen defense (Buer, Imin, & Djordjevic, 2010; Treutter, 2006). Lipids primarily constitute plant cell membranes, supply energy for germination and early growth, and play a role in stress-response signal transduction (Ohlrogge & Browse, 1995; Samuels, Kunst, & Jetter, 2008). In our study, content of most metabolite types in wild soybean seed coats was higher than that in the cultivars, and flavonoid quantities in embryos were greater in wild soybeans, suggesting a stronger environmental adaptability and potential nutritional substances of wild soybeans(Kozłowska & Szostak-Węgierek, 2017; Shomali et al., 2022). By identifying and reintroducing metabolites from wild soybeans, such as specific flavonoids and terpenoids that have been diminished or lost during the domestication process, advanced soybean varieties that not only meet agricultural needs but also address nutritional deficiencies prevalent in modern diets can be created.

The PLS-DA analysis revealed a distinction in the components of seed coats and embryos between the wild and cultivated soybean varieties (Fig. 2A and B). In the seed coat, components 1 and 2 explained 41.8% and 33.9% of the variance, respectively, totaling 75.7% of the variance. In embryos, the corresponding figures were 55.6% and 19.9%, respectively, totaling 75.5%. Variable important in projection (VIP) scoring, typically used to identify metabolites with the greatest differences in metabolomics analysis, highlighted the top 10 compounds in soybean seed coats as isovaleryl diethylamide, oleamide, xestoaminol C, gamma-aminobutyric acid, 13-docosenamide, lysophosphatidylcholine 18:2 (LPC(18:2)), 1S,3R-ACPD, *N*-acetyl-L-glutamic acid, L-arginine, and calendulaglycoside B, and in embryos as LPC(18:2), DL-arginine,
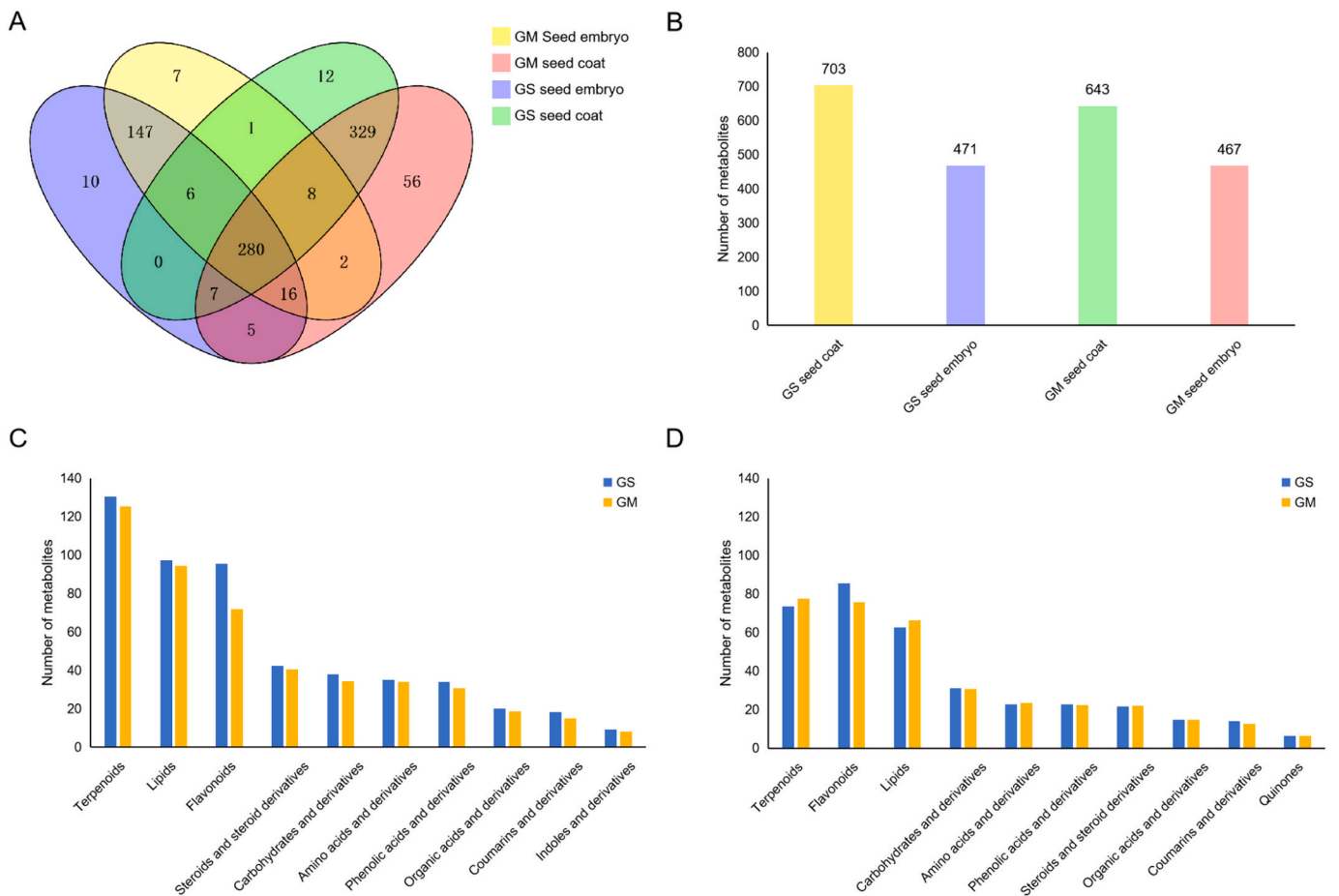
Fig. 1. Detection of metabolites in *G. soja* (GS-JL, GS-JS, GS-YN) and *G. max* (GM-HX10, GM-SD21, GM-ZLD) seed coats and embryos by LC-MS(+). A) Venn diagram. B) The number of metabolites. C) The number of top 10 metabolite types in soybean seed coats. D) The number top 10 metabolite types in soybean seed embryos.
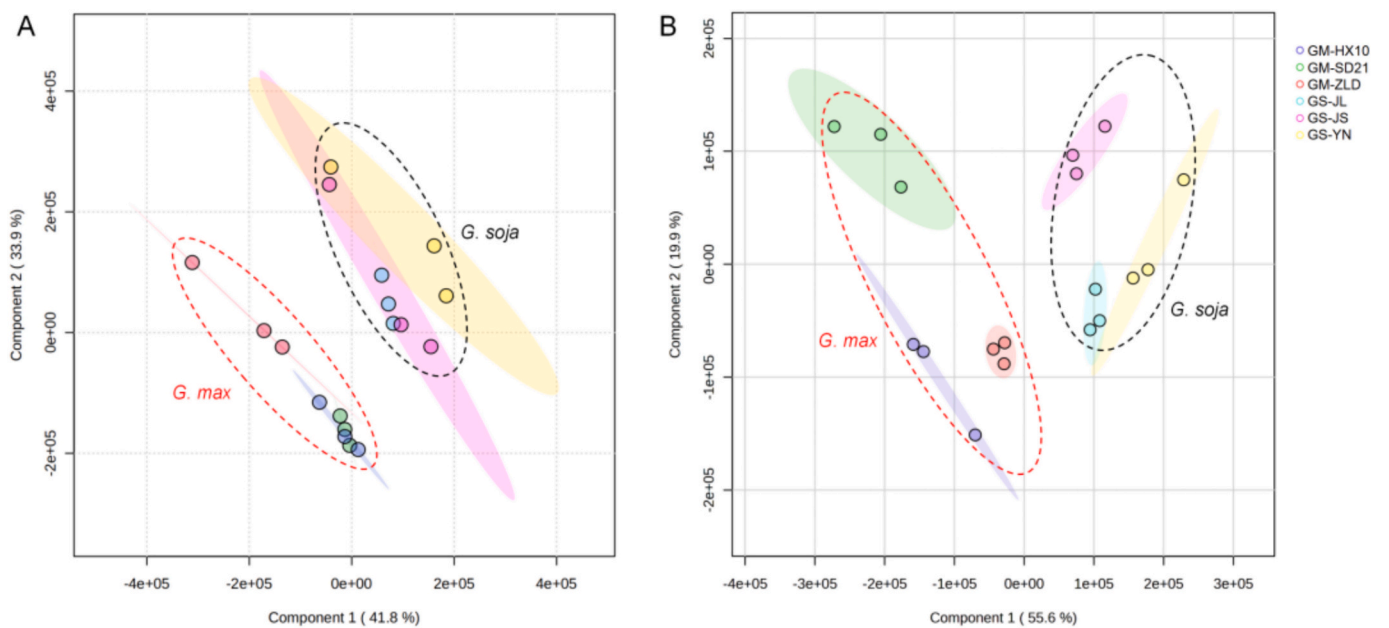
Fig. 2. PLS-DA 2D score plot of *G. soja* (GS-JL, GS-JS, GS-YN) and *G. max* (GM-HX10, GM-SD21, GM-ZLD). Ellipses represent 95% confidence interval (*n* = 3). A) Seed coats. B) Seed embryos.

palmitoylethanolamide (PEA), L-histidine, 3-oxoglutaric acid, 4-formyl-salicylic acid, L-asparagine, ethyl myristate, and 3-dehydrosphinganine (C20).

3.2. Differential metabolites in seed coats of both soybeans

Thirty-five differential metabolites in seed coats between wild and cultivated soybeans were identified, primarily including lipids, flavonoids, and amino acids and their derivatives using PLS-DA with VIP scores greater than or equal to 1. Of these, 21 metabolites showed higher expression in wild soybean seed coats than cultivars, such as malvidin-3-O-glucoside, quercetin 3-O-glucoside, and L-arginine, whereas 14 showed lower expression, including oleamide, gamma-aminobutyric acid, 13-docosenamide, and 1S,3R-ACPD (Fig. 3A, Table S2). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis revealed that the upregulated metabolites in wild soybean seed coats were mainly involved in aminoacyl-tRNA biosynthesis; arginine biosynthesis; nitrogen metabolism; valine, leucine, and isoleucine biosynthesis; and D-glutamine and D-glutamate metabolism. In contrast, the metabolites upregulated in cultivated soybeans participated in pathways such as alanine, aspartate, and glutamate metabolism; arginine biosynthesis; butanoate metabolism; and histidine metabolism (Fig. 3B). These identified flavonoids, lipids, and amino acids, known for their antioxidant and health-promoting properties, offer valuable insights for the fortification of food products and dietary supplements.

In this study, 12 lipid metabolites were identified in the seed coats of wild and cultivated soybeans, with higher expression levels of two prenol lipids (calendulaglycoside B and soybean saponin BG), both of which resist pathogens and pests (Muley, Khadabadi, & Banarase, 2009; B. Singh, Singh, Singh, & Kaur, 2017). Moreover, 1-octen-3-yl primeveroside (associated with the plant mechanical damage response) was detected (Alves et al., 2022). As previously discussed, flavonoids exhibit strong antioxidant, anti-inflammatory, and anti-allergic activities, play a crucial role in protecting plant cells from various stresses, and are integral to plant defense mechanisms (Husain & Mahmood, 2018; Ogura et al., 2016; Rüfer & Kulling, 2006). In this study, all four flavonoids

identified were upregulated in wild soybean seed coats, including quercetin 3-O-glucoside, (+)-catechin, glycitein, and procyanidin A2, which may be key metabolites contributing to the enhanced environmental adaptability of wild soybeans. In addition to flavonoids, abundant amino acid compounds such as L-arginine, isoleucine, and L-glutamine were identified in wild soybean seed coats, which play a significant role in nitrogen storage and utilization in wild soybean seeds. For instance, the alpha-amino group of glutamate transferred by various transaminases is crucial for synthesizing gamma-aminobutyric acid, arginine, and proline (Forde & Lea, 2007; Kobayashi, Kobayashi, Takahashi, Kumakura, & Matsuoka, 2021), and glutamate also functions in endogenous signaling pathways and plant responses to organic nitrogen (Lee, Liao, & Hsieh, 2023). The synthesis of L-arginine involves multiple metabolites and pathways, serving as a key means for nitrogen storage and transport (Winter, Todd, Trovato, Forlani, & Funck, 2015), and its metabolic pathway exhibits tissue specificity and is closely related to nitrogen metabolism in plants (Gaufichon et al., 2017). Additionally, higher levels of oleamide and linoleamide were detected in the seed coats of cultivated GM-ZLD soybean and some wild soybean individuals. These compounds inhibit the activity of SERCA-type Ca^{2+}-ATPases in a concentration-dependent manner, significantly regulating various physiological processes in plant cells, such as photosynthesis, growth, environmental stress response, and mechanical stimuli (Tanvir, Javeed, & Rehman, 2018). These key differential metabolites in seed coats contribute to understanding the environmental adaptability of wild soybeans and breeding efforts aimed at enhancing crop resilience and nutritional quality under varying climatic conditions, directly imoproving food quality. Notably, although terpenoids were the most abundant metabolites in both soybean seed coats, the differences between wild and cultivated soybeans were minor, and no differential metabolites were identified.

3.3. Differential metabolites in seed embryos of both soybeans

The difference in metabolites between the embryos of wild and cultivated soybeans was less pronounced compared with that in
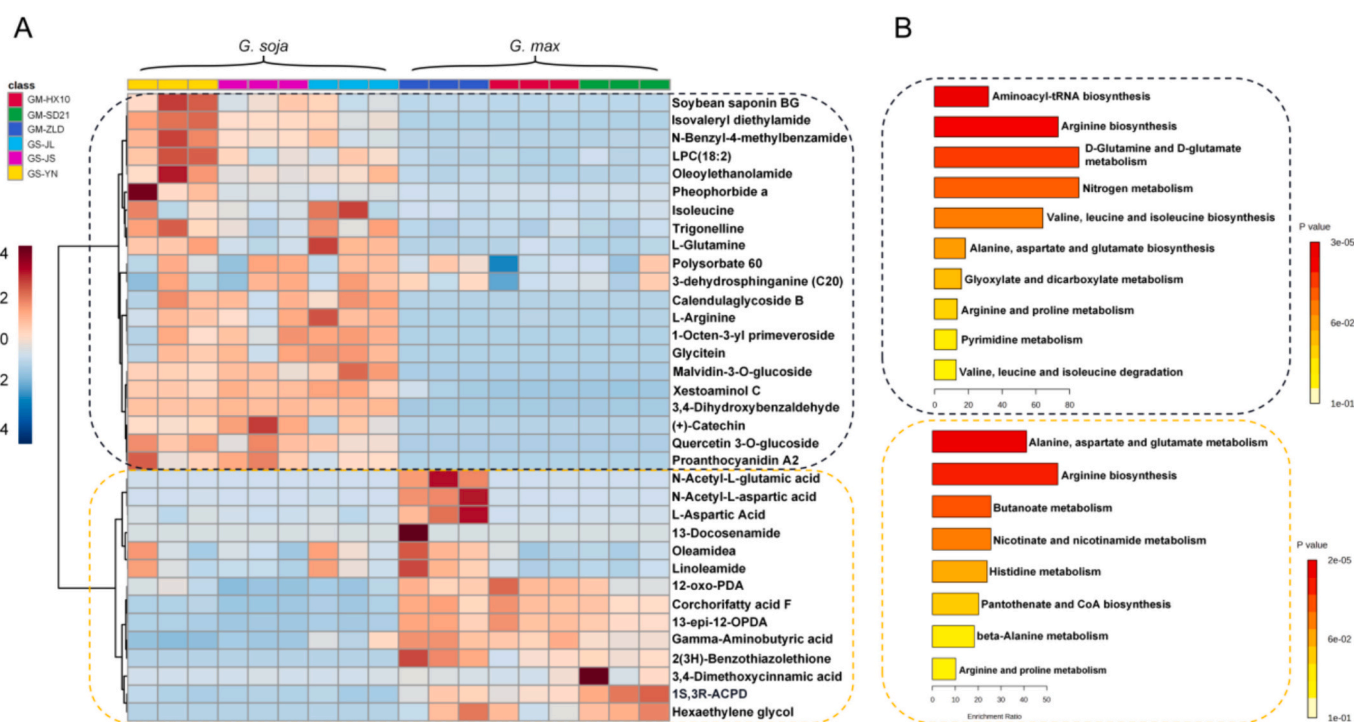

Fig. 3. Analysis of differential metabolites in seed coats between *G. soja* (GS-JL, GS-JS, GS-YN) and *G. max* (GM-HX10, GM-SD21, GM-ZLD). A) Pearson correlation cluster heatmap. B) KEGG annotation and enrichment.

metabolites between the seed coats, with only 10 differential metabolites identified (predominantly amino acids and lipids). Five metabolites were highly expressed in wild soybean embryos: PEA, L-asparagine, L-histidine, 4-formylsalicylic acid, and DL-arginine. Conversely, five metabolites were downregulated: LPC(18:2), ethyl myristate, cyanidin 3,5-diglucoside, 3-oxoglutaric acid, and 3-dehydrosphinganine(C20) (Table S3). Heatmaps of these differential metabolites showed significant variations between wild and cultivated soybeans (Fig. 4B). These metabolites are associated with aminoacyl-tRNA biosynthesis; histidine metabolism; beta-alanine metabolism; and alanine, aspartate, and glutamate metabolism according to KEGG enrichment analysis.

The five metabolites upregulated in wild soybean embryos are linked to plant defense and stress resistance. For instance, PEA (an endogenous agonist of peroxisome proliferator-activated receptor alpha (PPAR-α)) exhibits broad anti-inflammatory and antimicrobial activities. It also interacts with cannabinoid receptors, thereby reducing the production of inflammation-related reactive oxygen species (ROS) and nitric oxide (NO) (Morsanuto, Galla, Molinari, & Uberti, 2020). 4-Formylsalicylic acid is a derivative of salicylic acid and possesses salicylic acid functions, such as antioxidant properties, stress resistance, and roles in plant defense mechanisms and stress-related signaling pathways (Ding & Ding, 2020). L-histidine and L-asparagine are essential amino acids for plant growth, development, stress response, and nitrogen cycling (Seo et al., 2016). These metabolites may explain why wild soybeans have stronger environmental adaptability than cultivated soybeans.

3.4. In situ visualized distribution of differential metabolites

With the key differential metabolites identified, our focus shifted to understanding their spatial distribution within the seed structures of soybeans. Mapping these metabolites spatially provides crucial insights into their functional roles during the developmental stages and under various stress conditions, thus offering a more comprehensive view of their biological significance.

In this study, MALDI-MSI was used to construct ion images of 10 differential metabolites (VIP score > 1), visually illustrating their distribution in the embryos of wild and cultivated soybeans (Fig. 5). The spatial distribution of the differential metabolites of seed coats could not be demonstrated owing to their thin and cellulose-rich composition, which are difficult to ionize. In cultivated soybeans, 3-oxoglutaric acid was more concentrated and notably abundant in the hypocotyls and radicles, whereas it was evenly distributed across the cotyledons in wild soybeans. Although no studies have directly linked 3-oxoglutaric acid to the growth and development of soybean sprouts as a product and

substrate in a series of enzyme-catalyzed reactions within the tricarboxylic acid (TCA) cycle, it is crucial for providing energy (adenosine triphosphate, ATP) during soybean seed development and plays a significant role in various metabolic pathways, including amino acid synthesis and intracellular redox state regulation (Araújo, Martins, Fernie, & Tohge, 2014; Y. Zhang & Fernie, 2018). L-histidine and L-asparagine were uniformly distributed in wild and cultivated soybean embryos, with marked upregulation observed in the wild variety. Histidine may promote oil accumulation in seeds through ABA synthesis (Ma & Wang, 2016), and asparagine plays a vital role in regulating nitrogen metabolism during seed development (Credali et al., 2012), LPC(18:2) and cyanidin 3,5-diglucoside were significantly upregulated in cultivated soybeans, with a notably higher distribution in the cotyledons than in the hypocotyls and radicles. Lysophosphatidylcholines induce the production of ROS and ethylene, playing a critical role in plant defense responses (Wi, Seo, Cho, Nam, & Park, 2014). Anthocyanins affect seed color and exhibit important antioxidant activities (Kan, Nie, Hu, Liu, & Xie, 2016). DL-arginine was predominantly localized in the hypocotyl and radicle and was considerably upregulated in wild soybeans, with a notably high concentration in the GM-HX10 variety of cultivated soybeans. It promotes cell division in seeds, ensuring rapid growth of the soybean apical meristem for more efficient twinning and upward growth (Micallef & Shelp, 1989a, 1989b; Winter et al., 2015). Palmitoylethanolamide and 4-formylsalicylic acid were uniformly distributed and significantly more concentrated in wild soybeans, with PEA levels almost undetectable in the cultivated varieties. Conversely, ethyl myristate and 3-dehydrosphinganine(C20) showed a relatively uniform distribution in both species but were expressed at higher levels in cultivated soybeans.

3.5. Distribution of flavonoids in seed coats and embryos

In this study, flavonoids represent the class of compounds with the largest differences between wild and cultivated soybeans, and they also constitute the highest proportion among the differential metabolites identified. Furthermore, flavonoids not only have significant impacts on plant physiology but also offer numerous potential health benefits to humans. Research on flavonoid pathways at the molecular and cellular levels has contributed to our understanding of the environmental and developmental regulation of specialized metabolic pathways (Corso, Perreau, Mouille, & Lepiniec, 2020). Therefore, we further elucidated the distribution differences of flavonoids in the seed coats and embryos of wild and cultivated soybeans through visualization. There were significant differences in the accumulation and distribution of 15
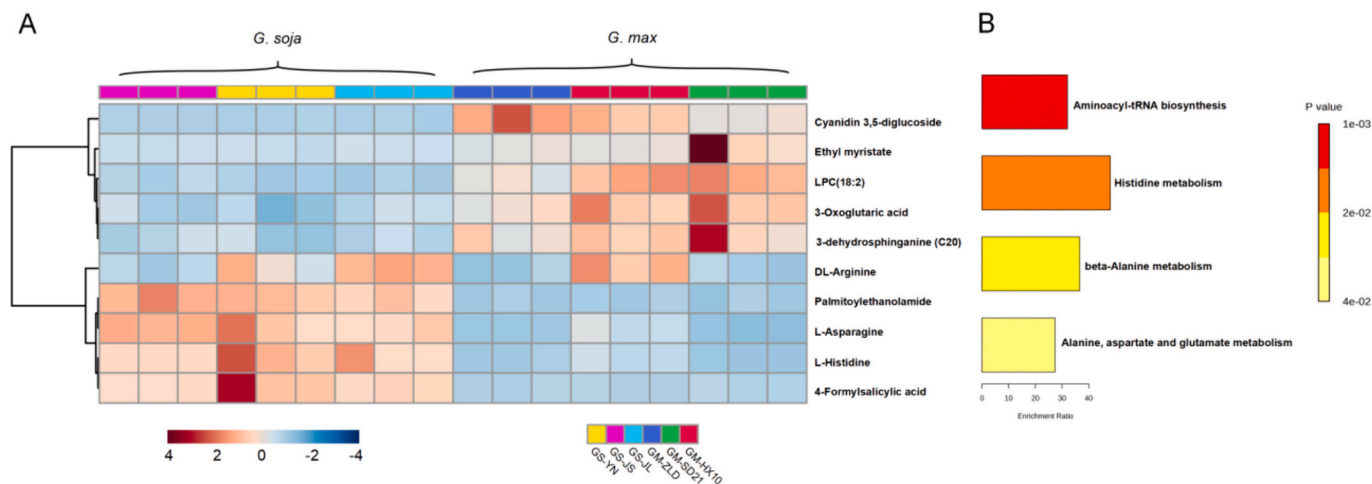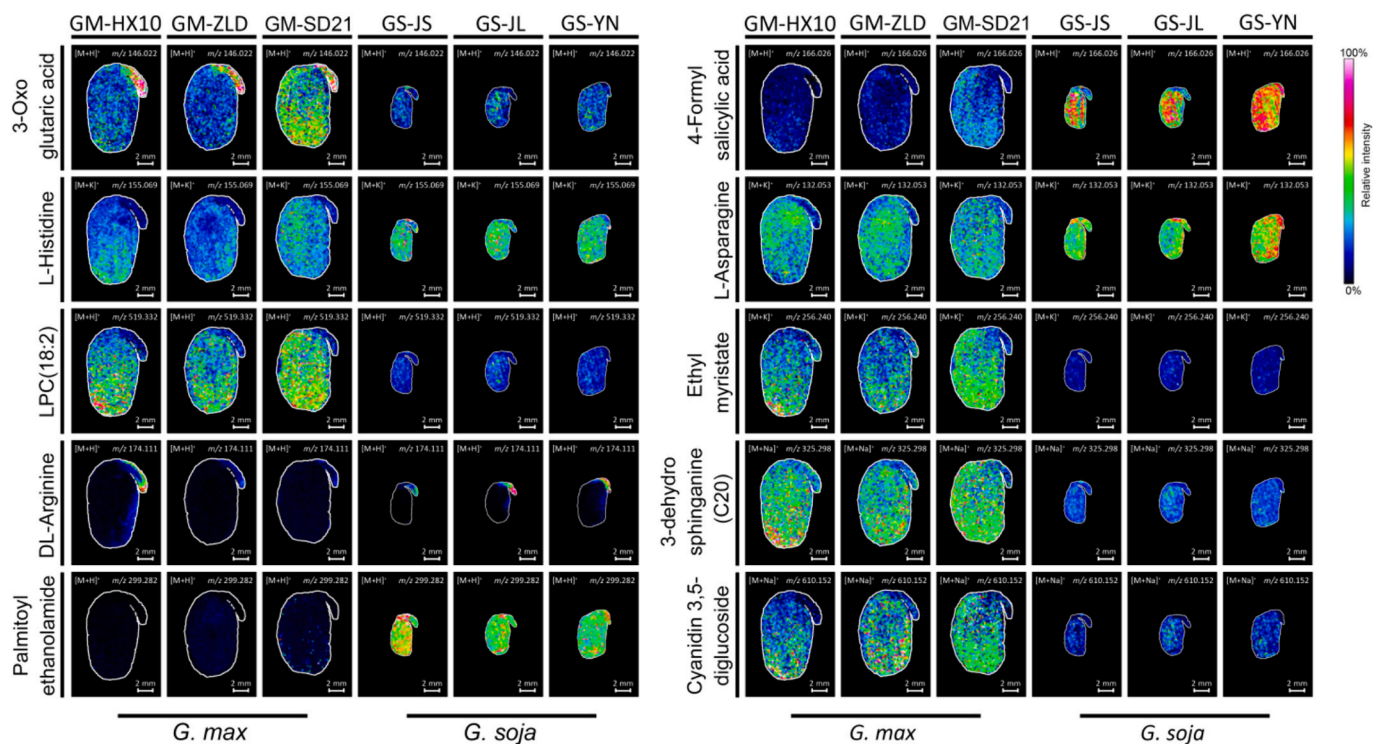

Fig. 4. Analysis of differential metabolites in seed embryos between *G. soja* (GS-JL, GS-JS, GS-YN) and *G. max* (GM-HX10, GM-SD21, GM-ZLD). A) Pearson correlation cluster heatmap. B) KEGG annotation and enrichment.

Fig. 5. Spatial distributions of differential metabolites between *G. soja* (GS-JL, GS-JS, GS-YN) and *G. max* (GM-HX10, GM-SD21, GM-ZLD) in embryos, visualized using MALDI-MSI.

flavonoids (isoliquiritigenin, 7,4′-dihydroxyflavone, phloretin, phlorizin, apigenin, luteolin, naringenin, dihydrokaempferol, kaempferol, (+)-afzelechin, quercetin, dihydromyricetin, (+)-catechin, (−)-epigallocatechin, and (−)-epicatechin) between the two species (Fig. 6). Notably, the levels of isoliquiritigenin, phlorizin, luteolin, and (+)-catechin were significantly higher in the seed coats and embryos of wild soybeans, influencing seed dormancy and germination by modulating plant hormone signaling and antioxidant activity (Ndakidemi & Dakora, 2003; Yilmaz & Toledo, 2004; S. Zhang et al., 2020). The distribution patterns of quercetin, (−)-epigallocatechin, and (−)-epicatechin were similar, predominantly located in the seed coats, with higher concentrations in wild soybeans, suggesting their role in preventing oxidative damage and pathogen resistance (Ahammed, Li, Li, Han, & Chen, 2018; Paszkowski & Kremer, 1988). Additionally, kaempferol and its precursors (naringenin and dihydrokaempferol) were mainly found in embryos, with kaempferol more abundant in cultivated soybean embryos, whereas no significant difference was observed for naringenin and dihydrokaempferol. Previous studies have suggested similar functions for flavonoids; however, the distinct distribution of different flavonoids in seed coats and embryos implies their potential sub-functionalities. For example, the higher content of quercetin in seed coats may play a more substantial role in pathogen defense and alleviating biotic and abiotic stresses during seed dormancy, while naringenin enrichment in embryos may influence seed germination and seedling rooting (Hernández & Munné-Bosch, 2012; P. Singh, Arif, Bajguz, & Hayat, 2021). Further in-depth research is necessary to reveal the synthesis, distribution, and functional differences of flavonoids in seeds.

4. Conclusions

In this study, a comparative metabolomic analysis of wild and cultivated soybeans was presented by utilizing MALDI-MSI and LC-MS technologies, revealing the inherent complexity and diversity of metabolic pathways in soybean seeds. A higher diversity and quantity of metabolites was identified, particularly terpenoids, flavonoids, and

lipids in wild soybeans, which are associated with growth, enhanced stress resistance and adaptability, such as L-histidine, PEA, LPC(18:2), and 4-Formylsalicylic acid. The unique metabolites in wild soybeans that were significantly reduced or lost in the cultivated varieties during domestication, like quercetin, luteolin, phlorizin, (+)-catechin, and (−)-epicatechin, suggesting the potential of wild soybeans as unique nutritional substances with higher ecological value.

The distinct metabolic characteristics of wild soybeans underscore the importance of preserving wild varieties. They are vital genetic resources for future breeding programs and may also harbor novel compounds that are beneficial for nutrition and health. Our findings highlight the potential of improving cultivated soybean varieties through the introduction of advantageous metabolic traits from wild soybean genetic resources. Future research should focus on the genetic pathways influencing these metabolite differences and the practical applications in crop breeding and biotechnology. Overall, the study highlights the crucial role of wild soybeans in agricultural sustainability and nutritional advancement, advocating for the continued protection and study of these invaluable genetic resources.

Author contributions

Xin Yin and Zhentao Ren wroted the manuscript. Ruizong Jia, Xiaodong Wang, Qi Yu, Li Zhang, Laipan Liu, Wenjing Shen, and Zhixiang Fang.contributed to the experiment and data analysis. Biao Liu and Jingang Liang designed the research and critically revised the manuscript. All authors read and approved the final manuscript.

CRediT authorship contribution statement

Xin Yin: Data curation, Formal analysis, Visualization, Writing – original draft. **Zhentao Ren:** Data curation, Formal analysis, Visualization, Writing – original draft. **Ruizong Jia:** Investigation, Resources. **Xiaodong Wang:** Funding acquisition, Investigation, Methodology. **Qi Yu:** Investigation, Resources. **Li Zhang:** Investigation, Resources.
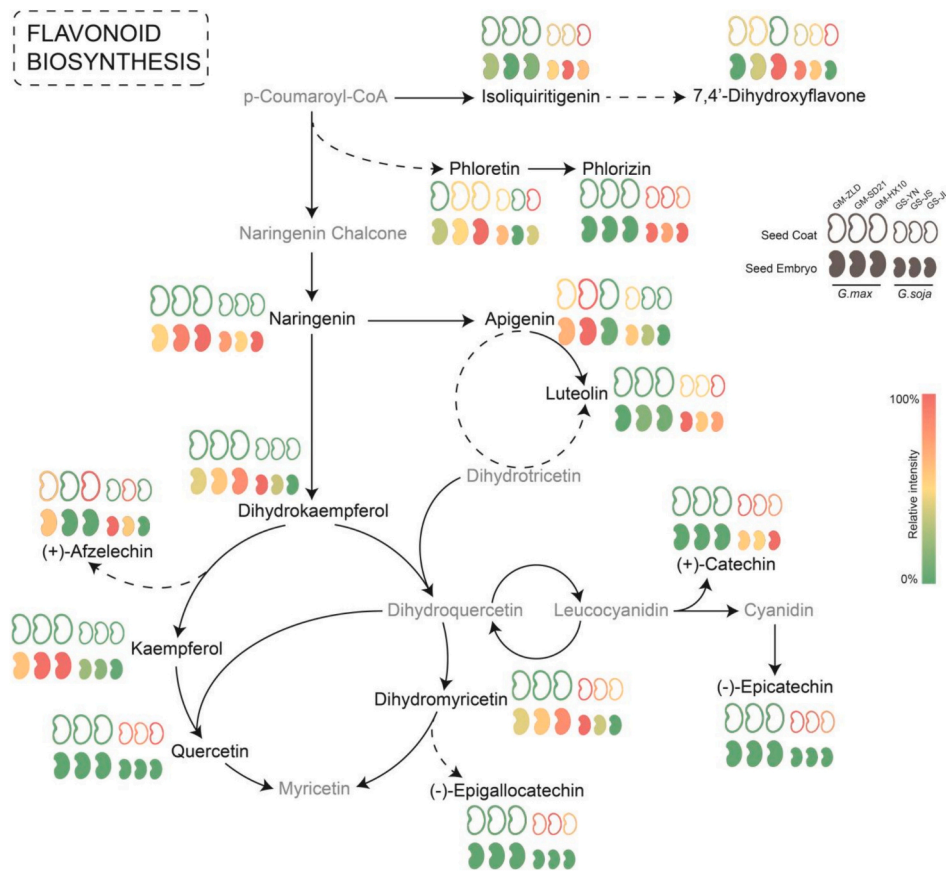
Fig. 6. Flavonoid biosynthesis pathway accumulation patterns in *G. soja* (GS-JL, GS-JS, GS-YN) and *G. max* (GM-HX10, GM-SD21, GM-ZLD) seed coats and embryos according to LC-MS (+). The hollow soybean represents the seed coat and solid soybean represents the embryos. The mean relative intensity (n = 3) is represented by color intensity, with red representing the highest and green the lowest concentrations. Compounds identified in this study are displayed in black. Solid lines indicate molecular interaction or relation, while dashed lines indicate indirect actions or unknown reaction. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Laipan Liu: Investigation, Resources. **Wenjing Shen:** Investigation, Resources. **Zhixiang Fang:** Investigation, Resources. **Jingang Liang:** Conceptualization, Writing – review & editing, Supervision. **Biao Liu:** Conceptualization, Funding acquisition, Project administration, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fochx.2024.101717.

References

Ahammed, G. J., Li, Y., Li, X., Han, W.-Y., & Chen, S. (2018). Epigallocatechin-3-gallate alleviates salinity-retarded seed germination and oxidative stress in tomato. *Journal of Plant Growth Regulation, 37*(4), 1349–1356. https://doi.org/10.1007/s00344-018-9849-0

Alves, R. J. M., Miranda, T. G., Pinheiro, R. O., de Souza Pinheiro, W. B., Andrade, E. H. d. A., & Tavares-Martins, A. C. C. (2022). Volatile chemical composition of *Octoblepharum albidum* Hedw. (Bryophyta) from the brazilian amazon. *BMC Chemistry, 16*(1), 76. https://doi.org/10.1186/s13065-022-00872-4

Anagnostopoulos, C., Stasinopoulou, P., Kanatas, P., & Travlos, I. (2020). Differences in metabolism of three *Conyza* species to herbicides glyphosate and triclopyr revealed by LC-MS/MS. *Chilean Journal of Agricultural Research, 80*, 100–107. https://doi.org/10.4067/S0718-58392020000100100

Araújo, W. L., Martins, A. O., Fernie, A. R., & Tohge, T. (2014). 2-oxoglutarate: Linking tca cycle function with amino acid, glucosinolate, flavonoid, alkaloid, and gibberellin biosynthesis. *Frontiers in Plant Science, 5*, 552. https://doi.org/10.3389/fpls.2014.00552

Bao, G., Mu, L., & Wang, Y. (2023). Effect of different accumulative temperate zones in Heilongjiang on *Glycine soja* metabolites as analyzed by non-target metabolomics. *Molecules, 28*, 3296. https://doi.org/10.3390/molecules28083296

Ben-Tov, D., Abraham, Y., Stav, S., Thompson, K., Loraine, A., Elbaum, R., Souza, A., Pauly, M., Kieber, J., & Harpaz-Saad, S. (2015). Cobra-like 2, a member of the gpi-anchored cobra-like family, plays a role in cellulose deposition in *Arabidopsis* seed coat mucilage secretory cells. *Plant Physiology, 167*(3), 711–724. https://doi.org/10.1104/pp.114.240671

Buer, C. S., Imin, N., & Djordjevic, M. A. (2010). Flavonoids: New roles for old molecules. *Journal of Integrative Plant Biology, 52*(1), 98–111. https://doi.org/10.1111/j.1744-7909.2010.00905.x

Cai, X., Jia, B., Sun, M., & Sun, X. (2022). Insights into the regulation of wild soybean tolerance to salt-alkaline stress. *Frontiers in Plant Science, 13*, 1002302. https://doi.org/10.3389/fpls.2022.1002302

Carter, T., Hymowitz, T., & Nelson, R. (2004). Biogeography, local adaptation, vavilov, and genetic diversity in soybean. In *Biological resources and migration* (pp. 47–59). Berlin: Springer.

Chen, J., Zhou, J., Li, M., Li, M., Hu, Y., Zhang, T., & Shi, L. (2022). Membrane lipid phosphorus reusing and antioxidant protecting played key roles in wild soybean resistance to phosphorus deficiency compared with cultivated soybean. *Plant and Soil, 474*, 99–113. https://doi.org/10.1007/s11104-022-05316-5

Chen, Y., & Nelson, R. L. (2004). Genetic variation and relationships among cultivated, wild, and semiwild soybean. *Crop Science, 44*(1), 316–325. https://doi.org/10.2135/cropsci2004.3160

Corso, M., Perreau, F., Mouille, G., & Lepiniec, L. (2020). Specialized phenolic compounds in seeds: Structures, functions, and regulations. *Plant Science, 296*, Article 110471. https://doi.org/10.1016/j.plantsci.2020.110471

Credali, A., García-Calderón, M., Dam, S., Perry, J., Díaz-Quintana, A., Parniske, M., … Márquez, A. J. (2012). The K+-dependent asparaginase, NSE1, is crucial for plant growth and seed production in *Lotus japonicus*. *Plant and Cell Physiology, 54*(1), 107–118. https://doi.org/10.1093/pcp/pcs156

Ding, P., & Ding, Y. (2020). Stories of salicylic acid: A plant defense hormone. *Trends in Plant Science, 25*(6), 549–565. https://doi.org/10.1016/j.tplants.2020.01.004

Dong, Y. S., Zhuang, B. C., Zhao, L. M., Sun, H., & He, M. Y. (2001). The genetic diversity of annual wild soybeans grown in China. *Theoretical and Applied Genetics, 103*(1), 98–103. https://doi.org/10.1007/s001220000522

Forde, B. G., & Lea, P. J. (2007). Glutamate in plants: Metabolism, regulation, and signalling. *Journal of Experimental Botany, 58*(9), 2339–2358. https://doi.org/10.1093/jxb/erm121

Fu, J., Gu, J., Bao, Z., Zhou, Y., Hu, H., Yang, C., Wu, R., Liu, H., Qin, L., Xu, H., Li, J., Guo, H., Wang, L., Zhou, Y., Wang, X., & Li, G. (2023). 2,5-dihydroxyterephthalic acid: A matrix for improved detection and imaging of amino acids. *Analytical Chemistry, 95*(51), 18709–18718. https://doi.org/10.1021/acs.analchem.3c01731

Gaufichon, L., Marmagne, A., Belcram, K., Yoneyama, T., Sakakibara, Y., Hase, T., Grandjean, O., Clément, G., Citerne, S., Boutet-Mercey, S., Masclaux-Daubresse, C., Chardon, F., Soulay, F., Xu, X., Trassaert, M., Shakiebaei, M., Najihi, A., & Suzuki, A. (2017). ASN1-encoded asparagine synthetase in floral organs contributes to nitrogen filling in *Arabidopsis* seeds. *Plant Journal, 91*(3), 371–393. https://doi.org/10.1111/tpj.13567

Gong, R., Huang, D., Chen, Y., Li, H., Wang, Z., Zhou, D., Zhao, L., Pan, Y., Chang, Y., Xiang, Y., Chongrong, W., & Zhou, S. (2020). Comparative metabolomics analysis reveals the variations of eating quality among three high-quality rice cultivars. *Molecular Breeding, 40*, 112. https://doi.org/10.1007/s11032-020-01192-y

Hernández, I., & Munné-Bosch, S. (2012). Naringenin inhibits seed germination and seedling root growth through a salicylic acid-independent mechanism in *Arabidopsis thaliana*. *Plant Physiology and Biochemistry, 61*, 24–28. https://doi.org/10.1016/j.plaphy.2012.09.003

Husain, N., & Mahmood, R. (2018). 3,4-dihydroxybenzaldehyde quenches ROS and RNS and protects human blood cells from Cr(VI)-induced cytotoxicity and genotoxicity. *Toxicology In Vitro, 50*, 293–304. https://doi.org/10.1016/j.tiv.2018.04.004

Hyeon, H., Xu, J., Kim, J. K., & Choi, Y. (2020). Comparative metabolic profiling of cultivated and wild black soybeans reveals distinct metabolic alterations associated with their domestication. *Food Research International, 134*, Article 109290. https://doi.org/10.1016/j.foodres.2020.109290

Kan, L., Nie, S., Hu, J., Liu, Z., & Xie, M. (2016). Antioxidant activities and anthocyanins composition of seed coats from twenty-six kidney bean cultivars. *Journal of Functional Foods, 26*, 622–631. https://doi.org/10.1016/j.jff.2016.08.030

Kim, M., Van, K., Kang, Y. J., Kim, K. H., & Lee, S.-H. (2012). Tracing soybean domestication history: From nucleotide to genome. *Breeding Science, 61*, 445–452. https://doi.org/10.1270/jsbbs.61.445

Kobayashi, W., Kobayashi, T., Takahashi, A., Kumakura, K., & Matsuoka, H. (2021). Metabolism of glutamic acid to alanine, proline, and γ-aminobutyric acid during takuan-zuke processing of radish root. *Journal of Food Science, 86*(2), 563–570. https://doi.org/10.1111/1750-3841.15567

Kofsky, J., Zhang, H., & Song, B.-H. (2018). The untapped genetic reservoir: The past, current, and future applications of the wild soybean (*Glycine soja*). *Frontiers in Plant Science, 9*, 949. https://doi.org/10.3389/fpls.2018.00949

Kozłowska, A., & Szostak-Węgierek, D. (2017). Plant flavonoids in health, prevention, and treatment of chronic diseases. In K. H. Al-Gubory, & I. Laher (Eds.), *Nutritional antioxidant therapies: Treatments and perspectives* (pp. 347–376). Cham: Springer International Publishing.

La, T., Large, E., Taliercio, E., Song, Q., Gillman, J. D., Xu, D., … Scaboo, A. (2019). Characterization of select wild soybean accessions in the USDA germplasm collection for seed composition and agronomic traits. *Crop Science, 59*(1), 233–251. https://doi.org/10.2135/cropsci2017.08.0514

Lafon-Placette, C., & Köhler, C. (2014). Embryo and endosperm, partners in seed development. *Current Opinion in Plant Biology, 17*, 64–69. https://doi.org/10.1016/j.pbi.2013.11.008

Langenheim, J. H. (1994). Higher plant terpenoids: A phytocentric overview of their ecological roles. *Journal of Chemical Ecology, 20*(6), 1223–1280. https://doi.org/10.1007/BF02059809

Leamy, L. J., Zhang, H., Li, C., Chen, C. Y., & Song, B.-H. (2017). A genome-wide association study of seed composition traits in wild soybean (*Glycine soja*). *BMC Genomics, 18*(1), 18. https://doi.org/10.1186/s12864-016-3397-4

Lee, K. T., Liao, H. S., & Hsieh, M. H. (2023). Glutamine metabolism, sensing, and signaling in plants. *Plant and Cell Physiology, 64*(12), 1466–1481. https://doi.org/10.1093/pcp/pcad054

Li, C.-R., Yang, L.-X., Guo, Z.-F., Yang, H., Zhang, Y., Wang, Y.-M., Zhang, G.-Z., Li, P., & Gao, W. (2022). LC-MS-based untargeted metabolomics reveals chemical differences of *Cannabis* leaves from different regions of China. *Industrial Crops and Products, 176*, Article 114411. https://doi.org/10.1016/j.indcrop.2021.114411

Li, Y.-H., Zhou, G., Ma, J., Jiang, W., Jin, L.-G., Zhang, Z., Guo, Y., Zhang, J., Sui, Y., Zheng, L., Zhang, S.-S., Zuo, Q., Shi, X.-H., Li, Y.-F., Zhang, W.-K., Hu, Y., Kong, G.,

Hong, H.-L., Tan, B., … Qiu, L.-J. (2014). *De novo* assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. *Nature Biotechnology, 32*(10), 1045–1052. https://doi.org/10.1038/nbt.2979

Liu, D., Li, M., Liu, Y., & Shi, L. (2020). Integration of the metabolome and transcriptome reveals the resistance mechanism to low nitrogen in wild soybean seedling roots. *Environmental and Experimental Botany, 175*, Article 104043. https://doi.org/10.1016/j.envexpbot.2020.104043

Llorach, R., Favari, C., Alonso, D., Garcia-Aloy, M., Andres-Lacueva, C., & Urpi-Sarda, M. (2019). Comparative metabolite fingerprinting of legumes using LC-MS-based untargeted metabolomics. *Food Research International, 126*, Article 108666. https://doi.org/10.1016/j.foodres.2019.108666

Ma, H., & Wang, S. (2016). Histidine regulates seed oil deposition through abscisic acid biosynthesis and β-oxidation. *Plant Physiology, 172*(2), 848–857. https://doi.org/10.1104/pp.16.00950

Micallef, B. J., & Shelp, B. J. (1989a). Arginine metabolism in developing soybean cotyledons : I. Relationship to nitrogen nutrition. *Plant Physiology, 90*(2), 624–630. https://doi.org/10.1104/pp.90.2.624

Micallef, B. J., & Shelp, B. J. (1989b). Arginine metabolism in developing soybean cotyledons: III. Utilization. *Plant Physiology, 91*(1), 170–174. https://doi.org/10.1104/pp.91.1.170

Morsanuto, V., Galla, R., Molinari, C., & Uberti, F. (2020). A new palmitoylethanolamide form combined with antioxidant molecules to improve its effectiveness on neuronal aging. *Brain Sciences, 10*(7), 457. https://doi.org/10.3390/brainsci10070457

Muley, B., Khadabadi, S., & Banarase, N. (2009). Phytochemical constituents and pharmacological activities of *Calendula officinalis* Linn (Asteraceae): A review. *Tropical Journal of Pharmaceutical Research, 8*(5), 455–465. https://doi.org/10.4314/tjpr.v8i5.48090

Nakayama, Y., & Yamaguchi, H. (2002). Natural hybridization in wild soybean (*Glycine max* ssp. *soja*) by pollen flow from cultivated soybean (*Glycine max* ssp. *max*) in a designed population. *Weed Biology and Management, 2*(1), 25–30. https://doi.org/10.1046/j.1445-6664.2002.00043.x

Ndakidemi, P. A., & Dakora, F. D. (2003). Legume seed flavonoids and nitrogenous metabolites as signals and protectants in early seedling development. *Functional Plant Biology, 30*(7), 729–745. https://doi.org/10.1071/fp03042

Ogura, K., Ogura, M., Shoji, T., Sato, Y., Tahara, Y., Yamano, G., … Nagashima, K. (2016). Oral administration of apple procyanidins ameliorates insulin resistance via suppression of pro-inflammatory cytokine expression in liver of diabetic Ob/Ob mice. *Journal of Agricultural and Food Chemistry, 64*(46), 8857–8865. https://doi.org/10.1021/acs.jafc.6b03424

Ohlrogge, J., & Browse, J. (1995). Lipid biosynthesis. *The Plant Cell, 7*(7), 957–970. https://doi.org/10.1105/tpc.7.7.957

Paszkowski, W. L., & Kremer, R. J. (1988). Biological activity and tentative identification of flavonoid components in velvetleaf (*Abutilon theophrasti* Medik.) seed coats. *Journal of Chemical Ecology, 14*(7), 1573–1582. https://doi.org/10.1007/BF01012523

Radchuk, V., & Borisjuk, L. (2014). Physical, metabolic and developmental functions of the seed coat. *Frontiers in Plant Science, 5*, 510. https://doi.org/10.3389/fpls.2014.00510

Rüfer, C. E., & Kulling, S. E. (2006). Antioxidant activity of isoflavones and their major metabolites using different in vitro assays. *Journal of Agricultural and Food Chemistry, 54*(8), 2926–2931. https://doi.org/10.1021/jf053112o

Samuels, L., Kunst, L., & Jetter, R. (2008). Sealing plant surfaces: Cuticular wax formation by epidermal cells. *Annual Review of Plant Biology, 59*(1), 683–707. https://doi.org/10.1146/annurev.arplant.59.103006.093219

Seo, S., Nakaho, K., Hong, S. W., Takahashi, H., Shigemori, H., & Mitsuhara, I. (2016). L-histidine induces resistance in plants to the bacterial pathogen *Ralstonia solanacearum* partially through the activation of ethylene signaling. *Plant and Cell Physiology, 57*(9), 1932–1942. https://doi.org/10.1093/pcp/pcw114

Shomali, A., Das, S., Arif, N., Sarraf, M., Zahra, N., Yadav, V., … Hasanuzzaman, M. (2022). Diverse physiological roles of flavonoids in plant environmental stress responses and tolerance. *Plants (Basel), 11*(22). https://doi.org/10.3390/plants11223158

Singh, B., Singh, J. P., Singh, N., & Kaur, A. (2017). Saponins in pulses and their health promoting activities: A review. *Food Chemistry, 233*, 540–549. https://doi.org/10.1016/j.foodchem.2017.04.161

Singh, P., Arif, Y., Bajguz, A., & Hayat, S. (2021). The role of quercetin in plants. *Plant Physiology and Biochemistry, 166*, 10–19. https://doi.org/10.1016/j.plaphy.2021.05.023

Tanvir, R., Javeed, A., & Rehman, Y. (2018). Fatty acids and their amide derivatives from endophytes: New therapeutic possibilities from a hidden source. *FEMS Microbiology Letters, 365*(12), Article fny114. https://doi.org/10.1093/femsle/fny114

Tareq, F., Kotha, R., Natarajan, S., Sun, J., & Luthria, D. (2023). An untargeted metabolomics approach to study the variation between wild and cultivated soybeans. *Molecules, 28*, 5507. https://doi.org/10.3390/molecules28145507

Tholl, D. (2015). Biosynthesis and biological functions of terpenoids in plants. In J. Schrader, & J. Bohlmann (Eds.), *Biotechnology of isoprenoids* (pp. 63–106). Cham: Springer International Publishing.

Treutter, D. (2006). Significance of flavonoids in plant resistance: A review. *Environmental Chemistry Letters, 4*(3), 147–157. https://doi.org/10.1007/s10311-006-0068-8

Wang, M., Li, W., Fang, C., Xu, F., Liu, Y., Wang, Z., Yang, R., Zhang, M., Liu, S., Lu, S., Lin, T., Tang, J., Wang, Y.-Q., Wang, H., Lin, H., Zhu, B., Chen, M., Kong, F., Liu, B., & Tian, Z. (2018). Parallel selection on a dormancy gene during domestication of crops from multiple families. *Nature Genetics, 50*, 1435–1441. https://doi.org/10.1038/s41588-018-0229-2

Wi, S. J., Seo, S., Cho, K., Nam, M. H., & Park, K. Y. (2014). Lysophosphatidylcholine enhances susceptibility in signaling pathway against pathogen infection through biphasic production of reactive oxygen species and ethylene in tobacco plants. *Phytochemistry, 104*, 48–59. https://doi.org/10.1016/j.phytochem.2014.04.009

Winter, G., Todd, C. D., Trovato, M., Forlani, G., & Funck, D. (2015). Physiological implications of arginine metabolism in plants. *Frontiers in Plant Science, 6*, 534. https://doi.org/10.3389/fpls.2015.00534

Wu, J., Cui, C., Zhao, H., Zhou, G., Qin, L., Li, X., Chen, L., Wang, X., & Wan, Y. (2022). In-situ detection and imaging of areca catechu fruit alkaloids by MALDI-MSI. *Industrial Crops and Products, 188*, Article 115533. https://doi.org/10.1016/j.indcrop.2022.115533

Xia, J., Sinelnikov, I., Han, B., & Wishart, D. (2015). Metaboanalyst 3.0—Making metabolomics more meaningful. *Nucleic Acids Research, 43*(W1), W251–W257. https://doi.org/10.1093/nar/gkv380

Yilmaz, Y., & Toledo, R. T. (2004). Major flavonoids in grape seeds and skins: Antioxidant capacity of catechin, epicatechin, and gallic acid. *Journal of Agricultural and Food Chemistry, 52*(2), 255–260. https://doi.org/10.1021/jf030117h

Zhang, H., Guang, C., Xu, H., Jing, S., Jiang, Y., Liu, Z., Zhang, H., Wang, F., Hu, X., & Zhu, Y. (2023). Transcriptome analysis of rice embryo and endosperm during seed germination. *International Journal of Molecular Sciences, 24*, 8710. https://doi.org/10.3390/ijms24108710

Zhang, S., Sun, S.-W., Shi, H.-L., Zhao, K., Wang, J., Liu, Y., Liu, X.-H., & Wang, W. (2020). Physiological and biochemical mechanisms mediated by allelochemical isoliquiritigenin on the growth of lettuce seedlings. *Plants, 9*(2), 245. https://doi.org/10.3390/plants9020245

Zhang, Y., & Fernie, A. R. (2018). On the role of the tricarboxylic acid cycle in plant productivity. *Journal of Integrative Plant Biology, 60*(12), 1199–1216. https://doi.org/10.1111/jipb.12690

Zhou, J., Li, Y., Wang, X., Liu, Y., David-Schwartz, R., Weissberg, M., Qiu, S., Guo, Z., & Yang, F. (2022). Analysis of *Elymus nutans* seed coat development elucidates the genetic basis of metabolome and transcriptome underlying seed coat permeability characteristics. *Frontiers in Plant Science, 13*, Article 970957. https://doi.org/10.3389/fpls.2022.970957