

Research Insight

Open Access

Key Genes Influencing Soybean Protein and Oil Content: Functional Insights

Qishan Chen ✉

Modern Agricultural Research Center, Cuixi Academy of Biotechnology, Zhuji, 311800, Zhejiang, China

✉ Corresponding email: qishan.chen@cuixi.orgBioscience Evidence, 2025, Vol.15, No.5 doi: [10.5376/be.2025.15.0024](https://doi.org/10.5376/be.2025.15.0024)

Received: 18 Aug., 2025

Accepted: 29 Sep., 2025

Published: 10 Oct., 2025

Copyright © 2025 Chen, This is an open access article published under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preferred citation for this article:

Chen Q.S., 2025, Key genes influencing soybean protein and oil content: functional insights, Bioscience Evidence, 15(5): 237-248 (doi: [10.5376/be.2025.15.0024](https://doi.org/10.5376/be.2025.15.0024))

Abstract This study mainly introduces the important genes related to protein and oil content in soybean seeds. In recent years, researchers have identified many major QTLs and candidate genes related to protein and oil content by using methods such as genome-wide association analysis, transcriptomics and proteomics. Especially on chromosomes 15 and 20, such as FAD2-1, GmSWEET10a/b, GmMFT, etc., these genes often affect both proteins and oils simultaneously, and often in the opposite direction. Many studies have also found that there is a significant negative correlation between proteins and oils, and their regulatory networks involve different pathways such as carbon metabolism, fatty acid synthesis, and sugar transport. In addition, some genes are also related to traits such as seed development and stress response, showing pleiotropy. The article also summarizes the functional verification of these genes and their application in molecular breeding. In the future, by integrating multi-omics data, machine learning and precision breeding technologies, it may help us break through the contradiction between protein and oil content, enhance the adaptability of soybeans, and meet different consumer demands. The purpose of this study is to provide a reference for the efficient improvement and sustainable utilization of soybeans.

Keywords Soybeans (*Glycine max*); Protein content; Oil content; Key genes; Molecular breeding

1 Introduction

Soybean (*Glycine max*) is one of the most important sources of protein and vegetable oil in the world. It is widely used in human food, animal feed and industrial raw materials. Soybean seeds contain approximately 40% protein and 20% oil on average, and play an important role in global nutrition security and agricultural development (Wang et al., 2021; Clevinger et al., 2023; Patel et al., 2025). With the increase in population and the improvement of living standards, the demand for high-protein and high-oil soybeans is growing larger and larger. But one problem is that there is often a negative correlation between protein and oil. When one component is increased, the other tends to decrease. This genetic contradiction has become an important obstacle to quality improvement (Patil et al., 2018; Lee et al., 2019; Wang et al., 2021; Kumar et al., 2022; Clevinger et al., 2023; Mo et al., 2024; Patel et al., 2025).

In recent years, the development of multi-omics technologies, such as genomics, transcriptomics and proteomics, has greatly promoted the research on the genetic basis of soy protein and oil. Through GWAS, QTL mapping and comprehensive analysis, researchers have identified many key genes and regulatory networks. For instance, genes such as GmSWEET10a/b, POWR1, GmMFT and FAD2-1B have been confirmed to play a core role in protein and oil accumulation. Some of these genes also exhibit pleiotropy, affecting not only proteins and oil content, but also seed size, yield and quality (Wang et al., 2020; Goettel et al., 2022; Kumar et al., 2022; Cai et al., 2023; Kim et al., 2023). In addition, some genes or QTLs have been found to weaken the negative correlation between proteins and oils, bringing new breakthrough opportunities for breeding (Li et al., 2018; Lee et al., 2019; Zhang et al., 2019; Kumar et al., 2022).

This study focuses on the genetic factors and molecular regulatory mechanisms of soy protein and oil content, as well as their applications in molecular breeding. The main objective is to summarize the key genes and their functions, sort out the molecular networks of protein and oil accumulation, analyze how basic research can be applied to breeding practice, and explore the prospects of molecular markers and gene editing in improving

soybean quality. It is hoped that by integrating the latest achievements, theoretical support and practical guidance can be provided for the cultivation of high-protein and high-oil content soybeans.

2 Overview of Soybean Protein and Oil Biosynthesis

2.1 Protein biosynthesis pathways: storage proteins (glycinin, β -conglycinin), regulatory nodes

There are mainly two types of proteins in soybean seeds: glycoprotein (glycinin, 11S) and β -conglycinin (β -conglycinin, 7S), which account for almost all of soy protein (Xu et al., 2022). Protein synthesis relies on amino acid supply, and amino acids come from metabolic processes such as glycolysis, the tricarboxylic acid cycle (TCA), and the pentose phosphate pathway (Xu et al., 2022; Mo et al., 2024). The key links include the synthesis, transportation and storage of amino acids, as well as the expression of protein genes. For example, the transcription factor GmESR1 can promote protein accumulation by regulating the phenylpropane pathway and sucrose transport. Sugar transporters (such as GmSWEET10a/b) also affect the distribution of proteins and fats (Jin et al., 2023).

2.2 Oil biosynthesis pathways: fatty acid and triacylglycerol synthesis, central metabolic regulation

Soybean oil mainly exists in the form of triacylglycerol (TAG). Fatty acid synthesis begins in plastids, and key enzymes include acetyl-coA carboxylase (ACCase), hydroxy-acyl-ACP dehydrase, and ketoacyl-ACP synthase, etc. (Xu et al., 2022; Zhao et al., 2023). The synthesized fatty acids are activated by long-chain acyl-coA synthase (LACS), then transported to the endoplasmic reticulum, and finally assembled into TAG (Xu et al., 2022; Zhao et al., 2023). Transcription factors such as GmWRI1a, GmVOZ1A, GmZF351, GmZF392, etc., can regulate related genes and significantly increase lipid content (Li et al., 2017; Chen et al., 2018; Lu et al., 2021; Yang et al., 2024; Wei et al., 2025). The DGAT (diacylglycerol acyltransferase) family genes play a rate-limiting role in TAG synthesis and are a key point for lipid accumulation (Zhao et al., 2023). Glycolysis and the TCA cycle are closely linked to fatty acid synthesis, and the distribution of carbon flow has an important impact on the accumulation of proteins and fats (Mo et al., 2024; Badia et al., 2025).

2.3 Nutritional and industrial significance: quality traits influencing end-use

Soy protein contains 18 kinds of amino acids, including all essential amino acids, and is an important source of nutrition for humans and animals. However, sulfur-containing amino acids (such as methionine and cysteine) are relatively scarce, which limits their nutritional value (Xu et al., 2022; Hooker et al., 2025). Soybean oil is mainly composed of linoleic acid, oleic acid and alpha-linolenic acid. The composition of fatty acids directly determines the nutrition and industrial quality of edible oil (Lakhssassi et al., 2021; Xu et al., 2022; Zhao et al., 2023). High-protein and high-oil content soybeans can not only increase the value of food, feed and industrial raw materials, but also meet the demands of healthy food, functional food and bioenergy (Lakhssassi et al., 2021; Xu et al., 2022). Therefore, understanding the molecular mechanisms of protein and lipid synthesis provides a foundation and direction for quality improvement and multi-purpose utilization.

3 Genetic Basis of Protein Content

3.1 Major QTLs and loci: insights from linkage mapping and GWAS

Soy protein content is a typical quantitative trait, which is jointly regulated by many genes. Through traditional linkage mapping and genome-wide association analysis (GWAS), researchers identified hundreds of related QTLs on the 20 chromosomes of soybeans. Especially on chromosomes 15 and 20, major QTLs (such as qPro15-1, qPro20-1, qSPC_20-1, qSPC_20-2) have been repeatedly located. These QTLs are stable in different populations and environments. It can explain more than 40% of the phenotypic differences (Van and McHale, 2017; Zhang et al., 2019; Wang et al., 2021; Zhang et al., 2022; Park et al., 2023). The latest GWAS has also discovered some new loci, such as the region near cBL-interacting protein kinases on chromosome 11 (Kim et al., 2023; Potapova et al., 2025). Meta-QTL analysis integrated over 80 studies and identified 55 protein-related Meta-Qtls, further narrowing the candidate range (Van and McHale, 2017).

3.2 Key genes:

3.2.1 Glyma. family genes regulating storage protein composition.

Some Glyma.family genes (such as Glyma.20G088000, Glyma.16G066600, Glyma.19G185700, Glyma.19G186000) have been repeatedly discovered within the principal QTL interval. They are directly involved in the synthesis of glycoproteins and β -concombulins (Wang et al., 2021; Park et al., 2023; Zhang et al., 2024). In addition, some new candidate genes (such as Glyma.11g015500, Glyma.20g050300) have also been confirmed to be related to protein content through GWAS (Kim et al., 2023).

3.2.2 Transcription factors (e.g., bZIP, NAC, MYB) influencing protein accumulation

Transcription factors are very important in the regulation of protein synthesis. Family members such as bZIP, NAC, and MYB can regulate storage protein genes, nitrogen metabolism, and seed enrichment, thereby affecting protein accumulation (Guo et al., 2022; Shen et al., 2022; Hooker et al., 2023; Tian et al., 2025). HSSP1 is a newly discovered regulatory factor. It can bind to the GmCG1 promoter, enhance its expression and increase seed protein. AIP2 affects the expression of glycoprotein and concombulin genes by regulating the ABI3 transcription factor.

3.3 Functional mechanisms: carbon–nitrogen balance, seed filling, transcriptional regulation

The accumulation of protein depends on the balance of carbon and nitrogen metabolism, amino acid synthesis and transport, as well as the distribution of substances during the seed enrichment stage. High-protein varieties usually have higher nitrogen utilization efficiency and stronger amino acid synthesis ability (Guo et al., 2022; Zhao et al., 2023; Zhang et al., 2024). Sugar transporters such as GmSWEET10a/b can affect carbon flow, thereby indirectly regulating protein and lipid accumulation (Wang et al., 2020). Furthermore, environmental factors, such as temperature and soil nitrogen, can also alter protein content by influencing gene expression (Hooker et al., 2023).

3.4 Breeding implications: challenges in raising protein without yield penalties

Although the discovery of major QTLs and key genes provides targets for molecular marker selection (MAS) and gene editing, protein content is often negatively correlated with yield and oil content. When protein is increased, it is often accompanied by a decrease in yield (Patil et al., 2017; Zhang et al., 2019; Wang et al., 2021; Guo et al., 2022; Liu et al., 2023). However, some specific QTLs (such as Chr.20 QTL derived from Danbaekkong) can increase the protein in specific genetic contexts without significantly reducing the yield (Patil et al., 2017; Park et al., 2023). In the future, it is necessary to combine multi-gene aggregation, precise editing and excellent germplasm resources to break the contradiction between protein and yield and promote the breeding of high-protein and high-yield soybeans.

4 Genetic Basis of Oil Content

4.1 Major QTLs and loci: landmark discoveries from genetic studies

Soybean seed oil content is a quantitative trait that is jointly controlled by many genes. Traditional linkage mapping and GWAS have identified hundreds of related QTLs on 20 chromosomes. Some QTLs on Chr.5, Chr.10, Chr.14 and Chr.20 (such as qOil-5-1, qOil-10-1, qOil-14-1, GqOil20) are very stable in different populations and environments. It can explain up to 26.3% of the phenotypic differences (Cao et al., 2017; Zhang et al., 2019; Jia et al., 2024). Meta-QTL and multi-omics analyses further narrowed the candidate regions and enhanced the value of breeding utilization (Jia et al., 2024; Yuan et al., 2024; Zhao et al., 2024).

4.2 Key genes

4.2.1 DGAT1 (diacylglycerol acyltransferase 1), WRI1, OLE1 regulators

DGAT1 is the rate-limiting enzyme for triacylglycerol (TAG) synthesis, and overexpression can significantly increase seed oil content (Zhao et al., 2024). WRI1 is a core transcription factor that can regulate genes related to glycolysis and fatty acid synthesis and promote lipid accumulation. GmOLEO1 encodes the structural proteins of oil bodies and affects the formation and stability of oil bodies. Its overexpression can increase oil content by more than 10% and was strongly selected during domestication (Zhang et al., 2019).

4.2.2 Acyl carrier proteins and desaturases (FAD2, FAD3)

Acyl carrier proteins (ACP) are responsible for transporting intermediates in fatty acid synthesis. FAD2 and FAD3 are key desaturases that respectively control the production of linoleic acid and alpha-linolenic acid, directly determine the ratio of oleic acid, linoleic acid, and alpha-linolenic acid, and are important targets for improving lipid quality (Yao et al., 2020; Ma et al., 2021; Silva et al., 2021; Cai et al., 2023). In addition, genes such as GmFATA1B and GmFATB1 also affect oil content and fatty acid composition (Ma et al., 2021; Cai et al., 2023).

4.3 Functional mechanisms: partitioning of carbon into lipid biosynthesis, transcriptional control

The synthesis of soybean oil relies on the distribution of carbon flow, shifting from precursors such as sucrose and glucose to fatty acids and TAG. Sugar transporters (such as GmSWEET10a/b, GmSWEET39) regulate oil accumulation by regulating the transport of sugar between seed coat and endosperm and affecting the utilization of carbon sources (Miao et al., 2019; Wang et al., 2020; 2025). Some transcription factors (such as WRI1, LEC1, MYB, bZIP, etc.) coordinate lipid synthesis by regulating genes related to fatty acid synthesis and TAG assembly (Zhang et al., 2019; Liu et al., 2020; Zhao et al., 2024).

4.4 Breeding implications: efforts to improve oil quality (oleic vs linoleic vs linolenic acid)

Increasing the oil content and optimizing the proportion of fatty acids (such as increasing oleic acid and reducing linoleic acid and alpha-linolenic acid) are important goals in soybean breeding. Through QTL mapping, molecular markers and gene editing, some excellent genotypes have been combined (Cao et al., 2017; Zhang et al., 2019; Ma et al., 2021; Silva et al., 2021; Cai et al., 2023; Jia et al., 2024). For example, the FAD2-1A/B mutant can stably increase the oleic acid content, and the knockout of GmFATB1 can reduce saturated fatty acids (Ma et al., 2021; Silva et al., 2021). However, there is a complex relationship between oil content, protein and yield. Multi-gene polymerization and precise regulation are required to achieve the breeding goals of high oil content and high quality (Zhang et al., 2019; Goettel et al., 2022; Jia et al., 2024; Yuan et al., 2024).

5 The Protein-Oil Trade-off

5.1 Physiological constraints: inverse relationship between protein and oil accumulation

In soybean seeds, the content of protein and oil is usually negatively correlated. That is to say, when protein levels increase, oil content tends to decrease, and vice versa (Wang et al., 2020; Liu et al., 2023). This relationship not only occurs among different varieties and germplasms, but is also obvious under different environments and management conditions of the same variety (Assefa et al., 2018; 2019). Genetic studies have shown that some major QTLS (such as loci on chromosomes 15 and 20) regulate proteins and oils simultaneously, and often in opposite directions (Zhang et al., 2020; Zhu et al., 2020; Clevinger et al., 2023; Kim et al., 2023). Environmental conditions also have an impact. For example, water stress, nitrogen fertilizer and irrigation can all change the accumulation of protein and oil. But overall, it is difficult to improve both simultaneously (Assefa et al., 2019; Carciochi et al., 2023).

5.2 Regulatory networks: hormonal and metabolic crosstalk

The relationship between protein and oil is not only determined by the genes themselves, but also related to the interaction of hormones and metabolic networks. Sugar transporters (such as GmSWEET10a/b, GmSWEET39) regulate the distribution of carbon sources such as sucrose between the seed coat and endosperm, thereby affecting whether carbon enters protein synthesis or lipid synthesis (Zhang et al., 2020; Wang et al., 2020; Mo et al., 2024). Some genes, such as GmMFT, can simultaneously affect seed development and material distribution, and have effects on both protein and oil (Cai et al., 2023). Furthermore, the interactive regulation of hormone signals (such as auxin and abscisic acid) and carbon-nitrogen metabolism is also regarded as an important mechanism of protein-oil balance (Duan et al., 2023; Mo et al., 2024) (Figure 1).

5.3 Systems biology perspectives: integrative omics evidence

Multi-omics studies (genomics, transcriptomics, proteomics, metabolomics) have shown that the accumulation of proteins and oils involves many metabolic pathways and regulatory links. Transcriptomic and proteomic analyses revealed differential expression of genes related to fatty acid and amino acid synthesis in high-protein or high-oil varieties. The changes in the activities of carbon metabolism, glycolysis and TCA cycle will also directly affect

the distribution of proteins and oils (Mo et al., 2024; Huang et al., 2025). Metabolomics studies further revealed that key metabolites such as glucose, citric acid, and α -ketoglutaric acid showed significant differences in content among different phenotypes, reflecting the competition of carbon sources between amino acid and fatty acid synthesis. Systems biology analysis has also identified a number of core genes that may simultaneously negatively regulate proteins and oils, providing new ideas for breaking through this trade-off (Kumar et al., 2021; Mo et al., 2024; Huang et al., 2025).

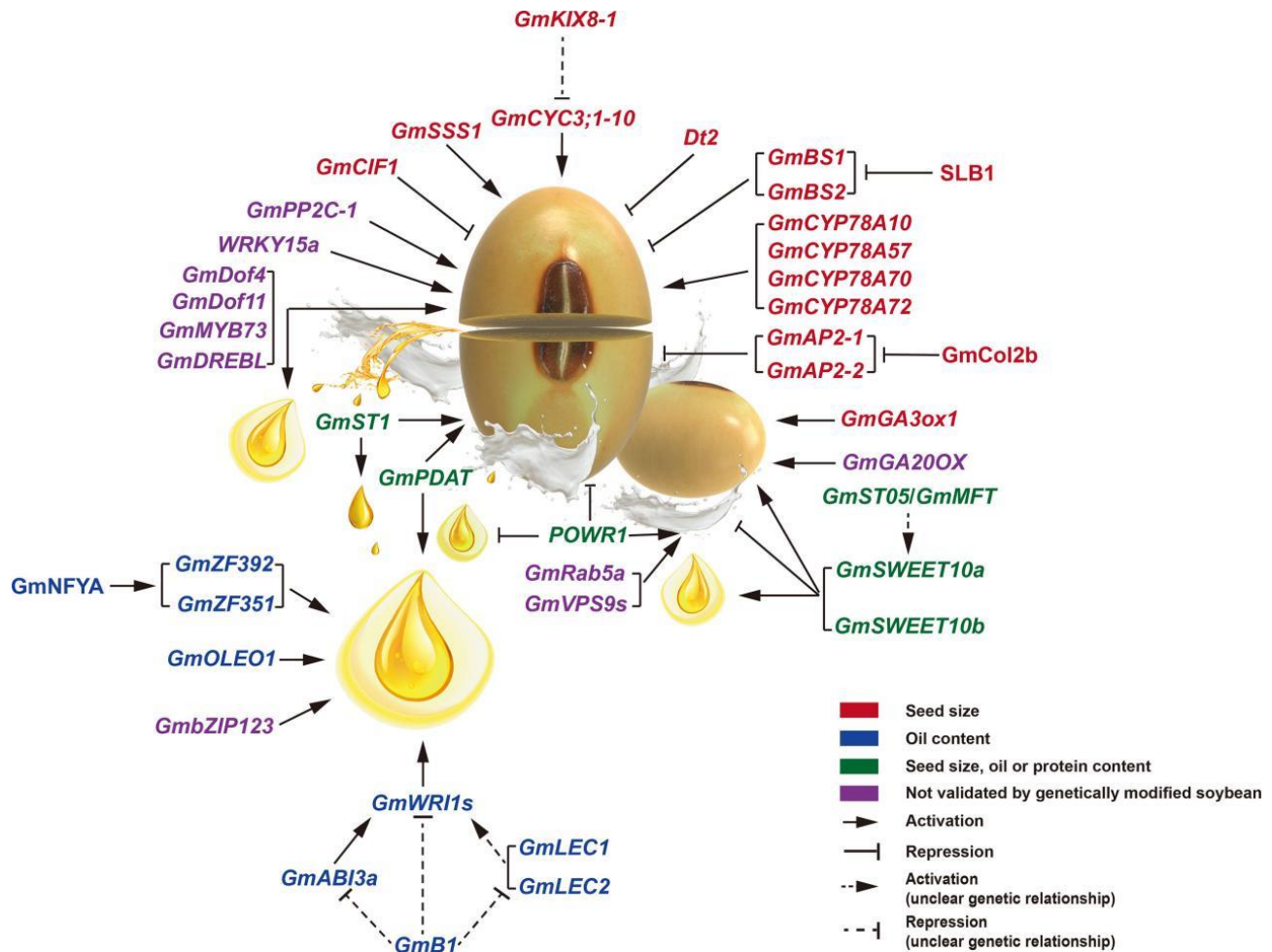


Figure 1 Genetic regulatory network of seed size (weight), oil accumulation, and protein content in soybean. The genes or proteins involving seed size (weight) and oil content are shown in red and blue fonts, respectively. The pleiotropic regulators for seed size (weight), oil accumulation, or protein content are indicated in green fonts. The regulatory genes, whose function has been validated only in Arabidopsis but not soybean, are shown in purple fonts (Adopted from Duan et al., 2023)

6 Advances in Biotechnological and Breeding Approaches

6.1 Conventional breeding: exploitation of natural variation

Conventional breeding mainly relies on methods such as hybridization, selection and mutagenesis, taking advantage of the natural variations in soybean germplasm resources to cultivate new varieties with high protein, high oil content and strong stress resistance. Both wild soybeans and local varieties worldwide offer rich genetic diversity (Anderson et al., 2019; Kumari et al., 2023; Vargas-Almendra et al., 2024). However, this method has a long cycle and limited genetic progression, making it difficult to meet the demand for rapidly increasing yield and quality (Anderson et al., 2019; Bhat and Yu, 2021).

6.2 Molecular breeding: marker-assisted selection, genomic selection

Molecular breeding has greatly enhanced the efficiency of improving complex traits. Marker-assisted selection (MAS) utilizes molecular markers closely linked to target traits to achieve precise aggregation of traits such as proteins, oils, and disease resistance (Bhat and Yu, 2021; Cao et al., 2022; Lin et al., 2022; Vargas-Almendra et al.,

2024) (Figure 2). Genomic selection (GS) uses genome-wide markers to predict breeding values and accelerate the screening speed of superior genotypes, and has been applied in disease resistance, stress resistance and quality improvement (Cao et al., 2022; Lin et al., 2022; Fang et al., 2023; Tian et al., 2025). Meanwhile, the combination of high-throughput genotyping and phenotypic techniques also enables molecular breeding to be faster and more accurate (Haidar et al., 2024).

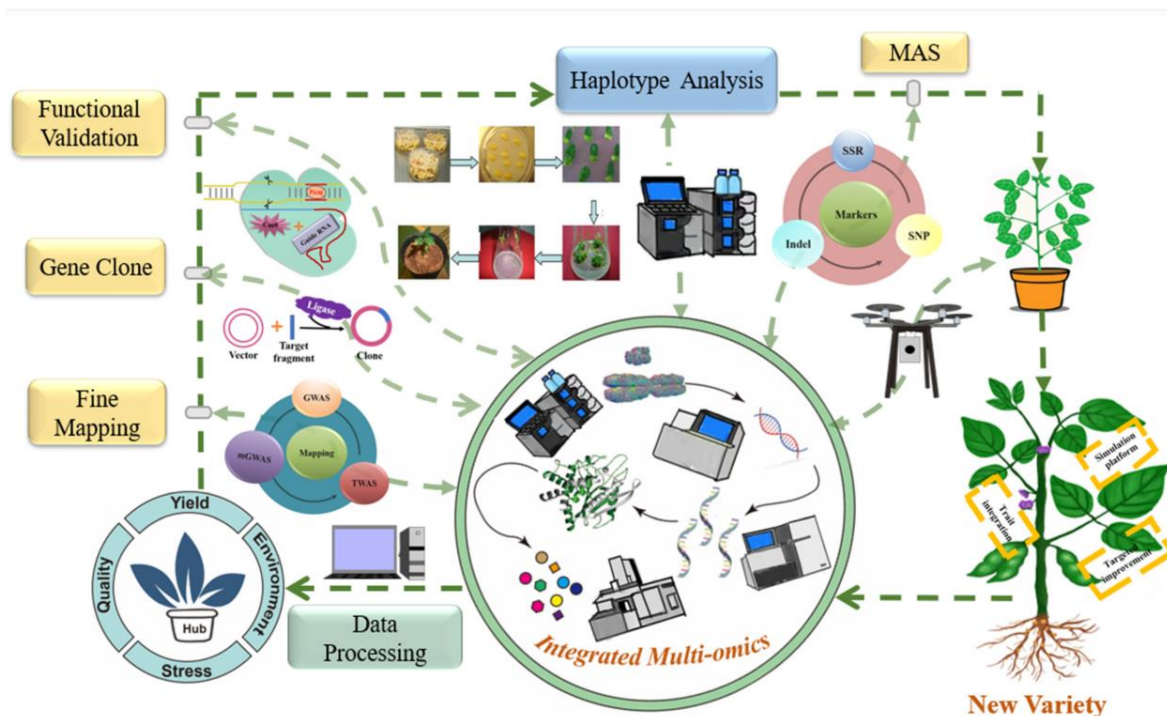


Figure 2 Multi-omics approaches for soybean molecular breeding (Adopted from Cao et al., 2022)

6.3 Genetic engineering: CRISPR/Cas9 examples, transgene approaches.

Genetic engineering offers new ways to regulate the content of proteins and oils. Editing techniques such as CRISPR/Cas9 can precisely knockout or modify key genes (such as FAD2, PDHK, etc.), and new soybean varieties with high oleic acid, low linolenic acid and high protein have been obtained (Gupta and Manjaya, 2022; Rahman et al., 2022; Xu et al., 2022; Monfort et al., 2025). Transgenic methods (such as Agrobacter mediation and gene gun method) are widely used for herbicide resistance, pest resistance, stress resistance and nutritional quality improvement. Currently, more than 80% of the global planting area is transgenic soybeans (Li et al., 2017; Anderson et al., 2019; Rahman et al., 2022; Xu et al., 2022). However, transgenic and gene editing still face challenges such as low transformation efficiency, genotype dependence and regulatory issues (Gai et al., 2025).

6.4 Synthetic biology: engineering carbon flux pathways.

Synthetic biology offers the possibility of simultaneously enhancing proteins and oils by modifying carbon metabolism pathways and optimizing the allocation of carbon and nitrogen. For instance, by regulating genes related to sugar transport, fatty acid synthesis and seed enrichment, the direction of carbon flow can be changed to enhance the nutritional and yield potential of seeds (Liu et al., 2020). In the future, by integrating multi-omics data and systems biology models, it is expected to promote the "design breeding" of soybean quality (Cao et al., 2022; Haidar et al., 2024; Tian et al., 2025).

7 Case Study: Functional Validation of Key Genes

7.1 Case study focus: example of a key gene such as GmWRI1a (oil regulator) or GmPDH1 (protein regulator)

7.1.1 Background: why this gene was prioritized

GmSWEET10a is a gene closely related to the quality of soybean seeds. It affects oil content, protein and grain size. Whole-genome resequencing and population analysis revealed that this gene was strongly selected during the

domestication of soybeans. Its different alleles are significantly related to the increase of oil content, the decrease of protein and the enlargement of grains. Therefore, it has become a key research object for functional verification and breeding (Wang et al., 2020).

7.1.2 Functional characterization: knockout/overexpression studies, phenotypic changes

Researchers used near-isogenic lines and genetically modified soybeans for verification. The results show that the superior alleles of GmSWEET10a can increase the oil content and grain weight of grains, but will reduce the protein content. It and the homologous gene GmSWEET10b are mainly responsible for transporting sucrose and hexose from the seed coat to the endosperm, regulating carbon source allocation, and thereby affecting lipid and protein accumulation (Wang et al., 2020). When overexpressed, the oil content of the grains was higher than that of the control. After knockout, the oil content decreased and the protein content increased.

7.1.3 Field validation: how findings translated into yield and composition shifts

In field trials, the oil content and yield of materials with superior alleles increased simultaneously, while the protein content slightly decreased. This effect is relatively stable in different environments and genetic backgrounds, indicating that it has application potential in production (Wang et al., 2020).

7.1.4 Breeding application: integration into elite lines

At present, the superior allele of GmSWEET10a has been introduced into modern major soybean varieties and has been used as a key target for MAS and gene editing. Meanwhile, its homologous gene GmSWEET10b has also begun to receive attention. In the future, by aggregating these superior genes, the oil content and yield of soybeans may be further increased (Wang et al., 2020).

7.2 Discussion: lessons learned—bridging functional genomics with practical breeding

The research case of GmSWEET10a shows that combining population genetics, transgenic or gene editing and field trials can achieve an effective transformation from gene discovery to breeding application. It not only explains the contradiction between protein and oil, but also provides a reference for the simultaneous improvement of multiple traits. More multi-omics data and precise editing methods are needed in the future to accelerate the application of functional genomics achievements in molecular breeding (Wang et al., 2020; Zhang et al., 2021; Kumar et al., 2022).

8. Future Directions

8.1 Integrative approaches: multi-omics, machine learning, and predictive breeding

In the future, the improvement of soy protein and oil content will rely more on the integration of multi-omics data (genomic, transcriptomic, proteomic, metabolomic). By integrating machine learning and big data analysis, complex traits can be predicted more accurately, enabling design breeding. Multi-omics studies have revealed the pathways of protein and oil accumulation, key genes and their interactions, which provide theoretical support for resolving the protein-oil contradiction (Kumar et al., 2021; Xu et al., 2022). Technologies such as machine learning and remote sensing have been used to efficiently predict protein and oil content in the field, improving the efficiency of breeding and field management (Hernandez et al., 2023). In the future, the combination of deep learning, data fusion and environmental factors will further enhance the accuracy of prediction and the ability of breeding decision-making (Patil et al., 2017).

8.2 Climate resilience: protein and oil traits under stress conditions.

Climate change poses new challenges to the protein and oil content of soybeans. Studies have shown that under future climatic conditions, oil content may increase with the increase in production, but protein content may decrease (Araji et al., 2020). The responses of different genotypes to stresses such as high temperature and drought are significantly different, which indicates that it is very important to screen and breed varieties with strong climate adaptability and stable quality (Araji et al., 2020; Duan et al., 2023). Multi-omics and QTL mapping can help identify genes related to stress response, providing a molecular basis for breeding new varieties with high protein, high oil content and stress resistance (Kumar et al., 2021; Xu et al., 2022; Duan et al., 2023).

8.3 Consumer-driven traits: high-protein varieties, healthy oil profiles

With people paying more attention to health and nutrition, the market demand for healthy soybean varieties such as high protein, high oleic acid, and low linolenic acid is constantly increasing (Song et al., 2023). High-protein varieties can not only meet the needs of the food and feed industries, but also increase the added value of soybeans (Guo et al., 2022). The improvement of lipid quality (such as increasing oleic acid and reducing saturated fatty acids) has become the focus of molecular breeding and gene editing (Song et al., 2023). In the future, by integrating molecular markers, gene editing and precise phenotypic screening, it is expected to achieve multi-objective synergistic improvement of protein, oil content and quality (Kumar et al., 2021).

9 Concluding Remarks

The protein and oil content of soybeans are very complex quantitative traits, controlled by many genes and QTLs together. Research has found that some major QTLs and candidate genes on chromosomes 15 and 20 (such as FAD2-1 and Glyma.20G085100) have been verified multiple times. They have significant effects on proteins and oils, and often in opposite directions. Multi-omics and proteomics studies also indicate that core metabolic pathways such as carbon metabolism, glycolysis, and the TCA cycle are of great significance, and related genes (such as sugar transporters, fatty acid synthases, signal transduction proteins, etc.) all play key roles in protein and oil accumulation. In addition, some genes also have pleiotropy, not only influencing seed components but also being related to stress response and adaptability.

A deep understanding of these genetic bases provides a foundation for modern breeding such as molecular marker-assisted selection, genomic selection and gene editing. The newly discovered QTL and functional genes provide molecular targets for cultivating soybeans with high protein, high oil content and high-quality fatty acid composition, which helps to meet the diverse needs of food, feed and industry. Meanwhile, some regulatory genes are related to stress tolerance, which provides new ideas for maintaining stable and sustainable soybean production under climate change.

However, the balance between protein and oil, environmental adaptability, and multi-objective collaborative improvement remain challenges in future research and breeding. In the future, it is necessary to integrate basic research achievements with new technologies such as molecular breeding and synthetic biology to accelerate the cultivation of high-quality, high-yield and sustainable new soybean varieties, and promote the global protein and oil industry towards green development.

Acknowledgments

The author appreciates the modification suggestions from Professors Rudi Mai and Qixue Liang on the manuscript of this study.

Conflict of Interest Disclosure

The authors affirm that this research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Anderson E., Ali L., Beavis W., Chen P., Clemente T., Diers B., Graef G., Grassini P., Hyten D., McHale L., Nelson R., Parrott W., Patil G., Stupar R., and Tilmon K., 2019, Soybean [*Glycine max* (L.) Merr.] breeding: history, improvement, production and future opportunities, *Advances in Plant Breeding Strategies: Legumes*, 431-516.
https://doi.org/10.1007/978-3-030-23400-3_12
- Araji H., Wayayok A., Daneshian J., Mirzaei M., Bavani A., Teh C., Abdullah A., and Ahmadi P., 2020, Effects of projected climate change on quantity and quality of soybean yield under different emission scenarios, *Current Science*, 118(1): 103-107.
<https://doi.org/10.18520/CS/V118/I1/103-107>
- Assefa Y., Bajjalieh N., Archontoulis S., Casteel S., Davidson D., Kovács P., Naeve S., and Ciampitti I., 2018, Spatial characterization of soybean yield and quality (amino acids, oil, and protein) for United States, *Scientific Reports*, 8: 14653.
<https://doi.org/10.1038/s41598-018-32895-0>
- Assefa Y., Purcell L., Salmerón M., Naeve S., Casteel S., Kovács P., Archontoulis S., Licht M., Below F., Kandel H., Lindsey L., Gaska J., Conley S., Shapiro C., Orlowski J., Golden B., Kaur G., Singh M., Thelen K., Laurenz R., Davidson D., and Ciampitti I., 2019, Assessing variation in US soybean seed composition (protein and oil), *Frontiers in Plant Science*, 10: 298.
<https://doi.org/10.3389/fpls.2019.00298>

- Badia M., Costa J., Zucchetti J., Pavlovic T., Calace P., Saigo M., and Wheeler M., 2025, Protein characterization of the soybean malic enzyme family to select metabolic targets for seed oil improvement, *bioRxiv*, 1-27.
<https://doi.org/10.1101/2025.03.18.644018>
- Bhat J., and Yu D., 2021, High-throughput NGS-based genotyping and phenotyping: role in genomics-assisted breeding for soybean improvement, *Legume Science*, 3(3): e81.
<https://doi.org/10.1002/leg3.81>
- Cai Z., Xian P., Cheng Y., Zhong Y., Yang Y., Zhou Q., Lian T., Ma Q., Nian H., and Ge L., 2023, MOTHER-OF-FT-AND-TFL1 regulates the seed oil and protein content in soybean, *The New phytologist*, 239(3): 905-919.
<https://doi.org/10.1111/nph.18792>
- Cao P., Zhao Y., Wu F., Xin D., Liu C., Wu X., Lv J., Chen Q., and Qi Z., 2022, Multi-omics techniques for soybean molecular breeding, *International Journal of Molecular Sciences*, 23(9): 4994.
<https://doi.org/10.3390/ijms23094994>
- Cao Y., Li S., Wang Z., Chang F., Kong J., Gai J., and Zhao T., 2017, Identification of major quantitative trait loci for seed oil content in soybeans by combining linkage and genome-wide association mapping, *Frontiers in Plant Science*, 8: 1222.
<https://doi.org/10.3389/fpls.2017.01222>
- Carciochi W., Grassini P., Naeve S., Specht J., Mamo M., Seymour R., Nygren A., Mueller N., Sivits S., Proctor C., Rees J., Whitney T., and La Menza N., 2023, Irrigation increases on-farm soybean yields in water-limited environments without a trade-off in seed protein concentration, *Field Crops Research*, 304: 109163.
<https://doi.org/10.1016/j.fcr.2023.109163>
- Chen L., Zheng Y., Dong Z., Meng F., Sun X., Fan X., Zhang Y., Wang M., and Wang S., 2018, Soybean (*Glycine max*) WRINKLED1 transcription factor, GmWRI1a, positively regulates seed oil accumulation, *Molecular Genetics and Genomics*, 293: 401-415.
<https://doi.org/10.1007/s00438-017-1393-2>
- Clevinger E., Biyashev R., Haak D., Song Q., Pilot G., and Maroof M., 2023, Identification of quantitative trait loci controlling soybean seed protein and oil content, *PLOS ONE*, 18(6): e0286329.
<https://doi.org/10.1371/journal.pone.0286329>
- Duan Z., Li Q., Wang H., He X., and Zhang M., 2023, Genetic regulatory networks of soybean seed size, oil and protein contents. *Frontiers in Plant Science*, 14: 1160418.
<https://doi.org/10.3389/fpls.2023.1160418>
- Fang X., Sun Y., Li J., Li M., and Zhang C., 2023, Male sterility and hybrid breeding in soybean, *Molecular Breeding: New Strategies in Plant Improvement*, 43: 47.
<https://doi.org/10.1007/s11032-023-01390-4>
- Gai Y., Liu S., Zhang Z., Wei J., Wang H., Liu L., Bai Q., Qin Q., Zhao C., Zhang S., Xiang N., and Zhang X., 2025, Integrative approaches to soybean resilience, productivity, and utility: a review of genomics, computational modeling, and economic viability, *Plants*, 14(5): 671.
<https://doi.org/10.3390/plants14050671>
- Goettel W., Zhang H., Li Y., Qiao Z., Jiang H., Hou D., Song Q., Pantalone V., Song B., Yu D., and An Y., 2022, POWR1 is a domestication gene pleiotropically regulating seed quality and yield in soybean, *Nature Communications*, 13: 3051.
<https://doi.org/10.1038/s41467-022-30314-7>
- Guo B., Sun L., Jiang S., Ren H., Sun R., Wei Z., Hong H., Luan X., Wang J., Wang X., Xu D., Li W., Guo C., and Qiu L., 2022, Soybean genetic resources contributing to sustainable protein production, *TAG. Theoretical and Applied Genetics*. Theoretische Und Angewandte Genetik, 135: 4095-4121.
<https://doi.org/10.1007/s00122-022-04222-9>
- Gupta S., and Manjaya J., 2022, Advances in improvement of soybean seed composition traits using genetic, genomic and biotechnological approaches, *Euphytica*, 218: 99.
<https://doi.org/10.1007/s10681-022-03046-4>
- Haidar S., Hooker J., Lackey S., Elian M., Puchacz N., Szczyglowski K., Marsolais F., Golshani A., Cober E., and Samanfar B., 2024, Harnessing multi-omics strategies and bioinformatics innovations for advancing soybean improvement: a comprehensive review, *Plants*, 13(19): 2714.
<https://doi.org/10.3390/plants13192714>
- Hernández C., Correndo A., Kyveryga P., Prestholt A., and Ciampitti I., 2023, On-farm soybean seed protein and oil prediction using satellite data, *Comput. Electron. Agric.*, 212: 108096.
<https://doi.org/10.1016/j.compag.2023.108096>
- Hooker J., Charette M., Lefebvre F., Zapata G., Mohr R., Daba K., Glenn A., Marsolais F., Hadinezhad M., Warkentin T., Hou A., Golshani A., Cober E., and Samanfar B., 2025, Differential expression analysis of soybean pod tissue between Canadian environments identifies differences in sulfur-containing amino acid-related gene expression, *Genome*, 68: 1-12.
<https://doi.org/10.1139/gen-2024-0106>
- Hooker J., Smith M., Zapata G., Charette M., Luckert D., Mohr R., Daba K., Warkentin T., Hadinezhad M., Barlow B., Hou A., Lefebvre F., Golshani A., Cober E., and Samanfar B., 2023, Differential gene expression provides leads to environmentally regulated soybean seed protein content, *Frontiers in Plant Science*, 14: 1260393.
<https://doi.org/10.3389/fpls.2023.1260393>

- Huang J., Ma Q., Cai Z., Xia Q., Li S., Jia J., Chu L., Lian T., Nian H., and Cheng Y., 2020, Identification and mapping of stable QTLs for seed oil and protein content in soybean [*Glycine max* (L.) Merr.], *Journal of Agricultural and Food Chemistry*, 68(23): 6448-6460.
<https://doi.org/10.1021/acs.jafc.0c01271>
- Huang Y., Meng B., Qin Y., Liu J., Lu A., Dai X., Zhao Y., and Ge L., 2025, Comparative proteomic atlas of two soybean varieties with contrasting seed oil and protein content, *Journal of Agricultural and Food Chemistry*, 73(4): 2279-2288.
<https://doi.org/10.1021/acs.jafc.4c07447>
- Jia H., Han D., Yan X., Zhang L., Liang J., and Lu W., 2024, Genome-wide association and RNA-Seq analyses reveal a potential candidate gene related to oil content in soybean seeds, *International Journal of Molecular Sciences*, 25(15): 8134.
<https://doi.org/10.3390/ijms25158134>
- Jin H., Yang X., Zhao H., Song X., Tsvetkov Y., Wu Y., Gao Q., Zhang R., and Zhang J., 2023, Genetic analysis of protein content and oil content in soybean by genome-wide association study, *Frontiers in Plant Science*, 14: 1182771.
<https://doi.org/10.3389/fpls.2023.1182771>
- Kim W., Kang B., Moon C., Kang S., Shin S., Chowdhury S., Choi M., Park S., Moon J., and Ha B., 2023, Quantitative trait loci (QTL) analysis of seed protein and oil content in wild soybean (*Glycine soja*), *International Journal of Molecular Sciences*, 24(4): 4077.
<https://doi.org/10.3390/ijms24044077>
- Kumar V., Goyal V., Mandlik R., Kumawat S., Sudhakaran S., Padalkar G., Rana N., Deshmukh R., Roy J., Sharma T., and Sonah H., 2022, Pinpointing genomic regions and candidate genes associated with seed oil and protein content in soybean through an integrative transcriptomic and QTL meta-analysis, *Cells*, 12(1): 97.
<https://doi.org/10.3390/cells12010097>
- Kumar V., Vats S., Kumawat S., Bisht A., Bhatt V., Shivaraj S., Padalkar G., Goyal V., Zargar S., Gupta S., Kumawat G., Chandra S., Chalam V., Ratnaparkhe M., Gill B., Jean M., Patil G., Vuong T., Rajcan I., Deshmukh R., Belzile F., Sharma T., Nguyen H., and Sonah H., 2021, Omics advances and integrative approaches for the simultaneous improvement of seed oil and protein content in soybean (*Glycine max* L.), *Critical Reviews in Plant Sciences*, 40: 398-421.
<https://doi.org/10.1080/07352689.2021.1954778>
- Kumari V., Thakur R., Kumari J., Kumari A., and Khajuria D., 2023, Nutritional improvement in soybean (*Glycine max* (L.) Merrill) through plant breeding and biotechnological interventions, *Crop and Pasture Science*, 75(1): CP23155.
<https://doi.org/10.1071/cp23155>
- Lakhssassi N., Zhou Z., Cullen M., Badad O., Baze A., Chetto O., Embaby M., Knizia D., Liu S., Neves L., and Meksem K., 2021, TILLING-by-sequencing+ to decipher oil biosynthesis pathway in soybeans: a new and effective platform for high-throughput gene functional analysis, *International Journal of Molecular Sciences*, 22(8): 4219.
<https://doi.org/10.3390/ijms22084219>
- Lee S., Van K., Sung M., Nelson R., LaMantia J., McHale L., and Mian M., 2019, Genome-wide association study of seed protein, oil and amino acid contents in soybean from maturity groups I to IV, *TAG. Theoretical and Applied Genetics*. *Theoretische Und Angewandte Genetik*, 132: 1639-1659.
<https://doi.org/10.1007/s00122-019-03304-5>
- Li S., Cong Y., Liu Y., Wang T., Shuai Q., Chen N., Gai J., and Li Y., 2017, Optimization of agrobacterium-mediated transformation in soybean, *Frontiers in Plant Science*, 8: 246.
<https://doi.org/10.3389/fpls.2017.00246>
- Li Y., Reif J., Hong H., Li H., Liu Z., Ma Y., Li J., Tian Y., Li Y., Li W., and Qiu L., 2018, Genome-wide association mapping of QTL underlying seed oil and protein contents of a diverse panel of soybean accessions, *Plant Science: An International Journal of Experimental Plant Biology*, 266: 95-101.
<https://doi.org/10.1016/j.plantsci.2017.04.013>
- Lin F., Chhapekar S., Vieira C., Da Silva M., Rojas A., Lee D., Liu N., Pardo E., Lee Y., Dong Z., Pinheiro J., Ploper L., Rupe J., Chen P., Wang D., and Nguyen H., 2022, Breeding for disease resistance in soybean: a global perspective, *TAG. Theoretical and Applied Genetics*. *Theoretische Und Angewandte Genetik*, 135: 3773-3872.
<https://doi.org/10.1007/s00122-022-04101-3>
- Liu S., Liu Z., Hou X., and Li X., 2023, Genetic mapping and functional genomics of soybean seed protein, *Molecular Breeding*, 43: 1-25.
<https://doi.org/10.1007/s11032-023-01373-5>
- Liu S., Zhang M., Feng F., and Tian Z., 2020, Toward a 'green revolution' for soybean, *Molecular Plant*, 13(5): 688-697.
<https://doi.org/10.1016/j.molp.2020.03.002>
- Lu L., Wei W., Li Q., Bian X., Lu X., Hu Y., Cheng T., Wang Z., Jin M., Tao J., Yin C., He S., Man W., Li W., Lai Y., Zhang W., Chen S., and Zhang J., 2021, A transcriptional regulatory module controls lipid accumulation in soybean, *The New Phytologist*, 231(2): 661-678.
<https://doi.org/10.1111/nph.17401>
- Ma J., Sun S., Whelan J., and Shou H., 2021, CRISPR/Cas9-mediated knockout of GmFATB1 significantly reduced the amount of saturated fatty acids in soybean seeds, *International Journal of Molecular Sciences*, 22(8): 3877.
<https://doi.org/10.3390/ijms22083877>
- Miao L., Yang S., Zhang K., He J., Wu C., Ren Y., Gai J., and Li Y., 2019, Natural variation and selection in GmSWEET39 affect soybean seed oil content, *The New Phytologist*, 225: 1651-1666.
<https://doi.org/10.1111/nph.16250>

- Mo W., Wang P., Shi Q., Zhao X., Zheng X., Ji L., Zhang L., Geng M., Wang Y., Wang R., Bian M., Meng X., Zuo Z., and Yang Z., 2024, Uncovering key genes associated with protein and oil in soybeans based on transcriptomics and proteomics, *Industrial Crops and Products*, 222(5): 119981.
<https://doi.org/10.1016/j.indcrop.2024.119981>
- Monfort M., Buitink J., Roeber F., and Nogué F., 2025, Genome editing, an opportunity to revive soybean cultivation in Europe, *The Plant Journal*, 121(4): e17266.
<https://doi.org/10.1111/tpj.17266>
- Park H., Seo J., Kang B., Kim J., Heo S., Choi M., Ko J., and Kim C., 2023, QTLs and candidate genes for seed protein content in two recombinant inbred line populations of soybean, *Plants*, 12(20): 3589.
<https://doi.org/10.3390/plants12203589>
- Patel J., Patel S., Cook L., Fallen B., and Koebernick J., 2025, Soybean genome-wide association study of seed weight, protein, and oil content in the southeastern USA, *Molecular Genetics and Genomics*, 300: 43.
<https://doi.org/10.1007/s00438-025-02228-8>
- Patil G., Vuong T., Kale S., Valliyodan B., Deshmukh R., Zhu C., Wu X., Bai Y., Yungbluth D., Lu F., Kumpatla S., Shannon J., Varshney R., and Nguyen H., 2018, Dissecting genomic hotspots underlying seed protein, oil, and sucrose content in an interspecific mapping population of soybean using high-density linkage mapping, *Plant Biotechnology Journal*, 16: 1939-1953.
<https://doi.org/10.1111/pbi.12929>
- Potapova N., Zorkoltseva I., Zlobin A., Shcherban A., Fedyaeva A., Salina E., Svishcheva G., Aksenovich T., and Tsepilov Y., 2025, Genome-wide association study on imputed genotypes of 180 Eurasian soybean *Glycine max* varieties for oil and protein contents in seeds, *Plants*, 14(2): 255.
<https://doi.org/10.3390/plants14020255>
- Rahman S., McCoy E., Raza G., Ali Z., Mansoor S., and Amin I., 2022, Improvement of soybean; a way forward transition from genetic engineering to new plant breeding technologies, *Molecular Biotechnology*, 65: 162-180.
<https://doi.org/10.1007/s12033-022-00456-6>
- Shen B., Schmidt M., Collet K., Liu Z., Coy M., Abbitt S., Molloy L., Frank M., Everard J., Booth R., Samadar P., He Y., Kinney A., and Herman E., 2022, RNAi and CRISPR-Cas silencing E3-ring ubiquitin ligase AIP2 enhances soybean seed protein content, *Journal of Experimental Botany*, 73(22): 7285-7297.
<https://doi.org/10.1093/jxb/erac376>
- Silva L., Da Matta L., Pereira G., Bueno R., Piovesan N., Cardinal A., God P., Ribeiro C., and Dal-Bianco M., 2021, Association studies and QTL mapping for soybean oil content and composition, *Euphytica*, 217: 24.
<https://doi.org/10.1007/s10681-020-02755-y>
- Song H., Taylor D., and Zhang M., 2023, Bioengineering of soybean oil and its impact on agronomic traits, *International Journal of Molecular Sciences*, 24(3): 2256.
<https://doi.org/10.3390/ijms24032256>
- Tian H., Yin Y., Li X., Zhang Z., Feng S., Jin S., Han X., Yang M., Xu C., Hu L., Liu C., Kong F., Chen Q., and Qi Z., 2025, Identification of HSSP1 as a regulator of soybean protein content through QTL analysis and Soy-SPCC network, *Plant Biotechnology Journal*, 23(7): 2673-2688.
<https://doi.org/10.1111/pbi.70092>
- Tian Z., Nepomuceno A., Song Q., Stupar R., Liu B., Kong F., Ma J., Lee S., and Jackson S., 2025, Soybean2035: a decadal vision for soybean functional genomics and breeding, *Molecular Plant*, 18(2): 245-271.
<https://doi.org/10.1016/j.molp.2025.01.004>
- Van K., and McHale L., 2017, Meta-analyses of QTLs associated with protein and oil contents and compositions in soybean [*Glycine max* (L.) Merr.] seed, *International Journal of Molecular Sciences*, 18(6): 1180.
<https://doi.org/10.3390/ijms18061180>
- Vargas-Almendra A., Ruiz-Medrano R., Núñez-Muñoz L., Ramírez-Pool J., Calderón-Pérez B., and Xoconostle-Cazares B., 2024, Advances in soybean genetic improvement, *Plants*, 13(21): 3073.
<https://doi.org/10.3390/plants13213073>
- Wang J., Mao L., Zeng Z., Yu X., Lian J., Feng J., Yang W., An J., Wu H., Zhang M., and Liu L., 2021, Genetic mapping high protein content QTL from soybean 'Nanxiadou 25' and candidate gene analysis, *BMC Plant Biology*, 21: 388.
<https://doi.org/10.1186/s12870-021-03176-2>
- Wang J., Zhang L., Wang S., Wang X., Li S., Gong P., Bai M., Paul A., Tvedt N., Ren H., Yang M., Zhang Z., Zhou S., Sun J., Wu X., Kuang H., Du Z., Dong Y., Shi X., Li M., Shukla D., Yan L., and Guan Y., 2025, AlphaFold-guided bespoke gene editing enhances field-grown soybean oil contents, *Advanced Science*, e2500290.
<https://doi.org/10.1002/advs.202500290>
- Wang S., Liu S., Wang J., Yokosho K., Zhou B., Yu Y., Liu Z., Frommer W., Ma J., Chen L., Guan Y., Shou H., and Tian Z., 2020, Simultaneous changes in seed size, oil content and protein content driven by selection of SWEET homologues during soybean domestication, *National Science Review*, 7: 1776-1786.
<https://doi.org/10.1093/nsr/nwaa110>
- Wei W., Wang L., Tao J., Zhang W., Chen S., Song Q., and Zhang J., 2025, The comprehensive regulatory network in seed oil biosynthesis, *Journal of Integrative Plant Biology*, 67(3): 649-668.
<https://doi.org/10.1111/jipb.13834>

- Xu H., Guo Y., Qiu L., and Ran Y., 2022, Progress in soybean genetic transformation over the last decade, *Frontiers in Plant Science*, 13: 900318.
<https://doi.org/10.3389/fpls.2022.900318>
- Yang M., Du C., Li M., Wang Y., Bao G., Huang J., Zhang Q., Zhang S., Xu P., Teng W., Li Q., Liu S., Song B., Yang Q., and Wang Z., 2024, The transcription factors GmVOZ1A and GmWRI1a synergistically regulate oil biosynthesis in soybean, *Plant Physiology*, 197(2): kiae485.
<https://doi.org/10.1093/plphys/kiae485>
- Yao Y., You Q., Duan G., Ren J., Chu S., Zhao J., Li X., Zhou X., and Jiao Y., 2020, Quantitative trait loci analysis of seed oil content and composition of wild and cultivated soybean, *BMC Plant Biology*, 20: 51.
<https://doi.org/10.1186/s12870-019-2199-7>
- Yuan X., Jiang X., Zhang M., Wang L., Jiao W., Chen H., Mao J., Ye W., and Song Q., 2024, Integrative omics analysis elucidates the genetic basis underlying seed weight and oil content in soybean, *The Plant Cell*, 36: 2160-2175.
<https://doi.org/10.1093/plcell/koae062>
- Zhang H., Goettel W., Song Q., Jiang H., Hu Z., Wan, M., and An Y., 2020, Selection of GmSWEET39 for oil and protein improvement in soybean, *PLoS Genetics*, 16(11): e1009114.
<https://doi.org/10.1371/journal.pgen.1009114>
- Zhang H., Zhang G., Zhang W., Wang Q., Xu W., Liu X., Cui X., Chen X., and Chen H., 2022, Identification of loci governing soybean seed protein content via genome-wide association study and selective signature analyses, *Frontiers in Plant Science*, 13: 1045953.
<https://doi.org/10.3389/fpls.2022.1045953>
- Zhang M., Liu S., Wang Z., Yuan Y., Zhang Z., Liang Q., Yang X., Duan Z., Liu Y., Kong F., Liu B., Ren B., and Tian Z., 2021, Progress in soybean functional genomics over the past decade, *Plant Biotechnology Journal*, 20: 256-282.
<https://doi.org/10.1111/pbi.13682>
- Zhang T., Wu T., Wang L., Jiang B., Zhen C., Yuan S., Hou W., Wu C., Han T., and Sun S., 2019, A combined linkage and GWAS analysis identifies QTLs linked to soybean seed protein and oil content, *International Journal of Molecular Sciences*, 20(23): 5915.
<https://doi.org/10.3390/ijms20235915>
- Zhang X., Wang F., Chen Q., Zhao Q., Zhao T., Hu X., Liu L., Qi J., QiaomY., Zhang M., Yang C., and Qin J., 2024, Identification of QTLs and candidate genes for water-soluble protein content in soybean seeds, *BMC Genomics*, 25: 783.
<https://doi.org/10.1186/s12864-024-10563-0>
- Zhao Q., Ma N., Li R., Zhong Y., Li X., and Liao H., 2023, Seed protein genetics linked with nitrogen and phosphorus translocation efficiency in soybean, *Agronomy*, 13(2): 598.
<https://doi.org/10.3390/agronomy13020598>
- Zhao X., Zhang Y., Wang J., Zhao X., Li Y., Teng W., Han Y., and Zhan Y., 2024, GWAS and WGCNA analysis uncover candidate genes associated with oil content in soybean, *Plants*, 13(10): 1351.
<https://doi.org/10.3390/plants13101351>
- Zhu X., Leiser W., Hahn V., and Würschum T., 2020, Identification of seed protein and oil related QTL in 944 RILs from a diallel of early-maturing European soybean, *Crop Journal*, 9(1): 238-247.
<https://doi.org/10.1016/j.cj.2020.06.006>



Disclaimer/Publisher's Note

The statements, opinions, and data contained in all publications are solely those of the individual authors and contributors and do not represent the views of the publishing house and/or its editors. The publisher and/or its editors disclaim all responsibility for any harm or damage to persons or property that may result from the application of ideas, methods, instructions, or products discussed in the content. Publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
