

randomly; rather, they reflect tightly coordinated developmental programs involving carbon and nitrogen allocation, fatty acid and triacylglycerol biosynthesis, and synthesis of major storage proteins such as β -conglycinin and glycinin. At the same time, oil and protein accumulation are linked by a persistent inverse relationship, indicating competition for shared metabolic precursors during seed development (Niu et al., 2025). Understanding when and how this trade-off is established is therefore essential for clarifying seed developmental physiology and for identifying routes to improve both traits simultaneously.

Research on soybean seed composition dynamics has shown that protein and oil contents change substantially over the course of development rather than appearing only as fixed mature-seed traits (Kambhampati et al., 2021). Early work established that oil percentage rises rapidly during mid-development, while large proportions of total mature protein and oil are synthesized during later filling stages even after percentage values begin to stabilize. Subsequent studies refined this developmental picture by showing that protein content often declines during the first weeks after flowering and then increases gradually, whereas oil accumulates rapidly at earlier stages, and sugars and oligosaccharides follow distinct temporal patterns toward maturity. Stage-based analyses further demonstrated that between reproductive stages R5 and R7, most mature dry weight is accumulated together with marked increases in protein, oil, and sugars, while after R7 moisture declines rapidly and stored-component composition changes relatively little. Collectively, these findings indicate that developmental timing is a major determinant of final seed composition and that the kinetics of reserve deposition must be considered explicitly in studies of soybean quality formation.

Despite substantial progress, important knowledge gaps remain in understanding the regulatory basis of dynamic protein and oil accumulation in soybean seeds. Recent genomic and transcriptomic studies have greatly expanded knowledge of the enzymes, transcription factors, and co-expression modules associated with reserve biosynthesis, especially during late developmental stages when divergence between high-oil and high-protein genotypes becomes most pronounced. Evidence now suggests that late maturation involves antagonistic activation of lipid-centered and nitrogen-centered pathways, providing a systems-level explanation for the oil-protein trade-off (Niu et al., 2025). However, many genetically mapped loci remain functionally unresolved, only a small number of candidate genes have been validated, and technical limitations in soybean transformation continue to slow mechanistic testing of proposed regulators (Duan et al., 2023). In addition, metabolic studies show that mature-seed composition can obscure important temporal shifts, including lipid decline and carbohydrate redistribution during maturation, meaning that endpoint measurements alone are insufficient to explain how seed composition is formed (Kambhampati et al., 2021). Accordingly, a focused analysis of dynamic changes in protein and oil contents during soybean seed development remains necessary to connect developmental stage, metabolic flux, and regulatory control, and to support breeding strategies aimed at improving seed quality without reinforcing the traditional trade-off between these two economically critical reserves.

2 Developmental Stages of Soybean Seeds and Physiological Regulation of Storage Compound Accumulation

2.1 Morphological and physiological characteristics during soybean seed development

Soybean seed development is commonly divided into lag, seed-filling, and maturation phases, and these phases can also be resolved morphologically from cotyledon formation through early maturity, mid-maturity, late maturity, and dry seed stages. Across these stages, seeds enlarge progressively, attain maximum size before desiccation, and then lose water rapidly as they approach physiological maturity, making morphology and water status reliable indicators of developmental progression. Physiologically, the major increase in seed mass and reserve deposition occurs mainly between R5 and R7, when moisture declines gradually but dry weight, protein, oil, and sugars all rise sharply. Cell division is largely completed by R4, whereas the pronounced increase in seed size from R5 to R6 is driven primarily by cell enlargement, during which carbon is partitioned simultaneously into protein, oil, and carbohydrate reserves (Islam et al., 2021).