

accumulation program during seed maturation (Jo et al., 2024). Consistent with this framework, transcriptomic analyses of developing seeds show that rapid oil accumulation is accompanied by coordinated upregulation of carbon fixation, photosynthesis, glycolysis, and fatty acid biosynthesis pathways, indicating that lipid deposition depends on integrated activation of precursor supply and biosynthetic capacity.

Beyond the LEC1-WRI1 axis, soybean oil accumulation is fine-tuned by additional transcription factors and downstream metabolic enzymes. GmZF392 functions as a positive regulator of lipid production by activating lipid biosynthesis genes, acts synergistically with GmZF351, and is positioned downstream of GmNFYA within a three-factor regulatory module that enhances seed oil accumulation (Lu et al., 2021). A second layer of control involves protein-protein cooperation, as GmVOZ1A interacts with GmWRI1a, jointly upregulates GmACBP6a, and thereby promotes TAG accumulation, while broader network analyses in cultivated soybean also identify a BCCP2-SAD-FAD2-OBO/FA9 axis and a PLIP1-dependent pathway linked to enhanced oil biosynthesis (Yang et al., 2024; Niu et al., 2026).

3.3 Coordination between carbon and nitrogen metabolism

The balance between protein and oil accumulation in soybean seeds is shaped by competition for shared carbon skeletons and by developmental regulation of nitrogen assimilation. Sucrose imported into developing seeds is metabolized through glycolysis to produce intermediates such as acetyl-CoA and phosphoenolpyruvate, which support both fatty acid and amino acid biosynthesis, making carbon allocation a central determinant of storage compound partitioning (Qi et al., 2026). This shared metabolic dependence helps explain the widely observed inverse relationship between seed oil and protein, which reflects repartitioning among protein, lipid, and carbohydrate reserves rather than simple independent accumulation of each component (Qi et al., 2026).

Evidence from multi-omics, physiology, and transporter genetics shows that this trade-off is dynamically regulated late in seed development. High-oil and high-protein cultivars diverge mainly during later maturation, when lipid-centric pathways such as TAG synthesis and oil body biogenesis are antagonistically activated against nitrogen-centric pathways including nitrogen assimilation, amino acid metabolism, ABA signaling, and storage protein processing (Niu et al., 2025). Experimental manipulation of maternal carbon supply supports this model: reduced sugar delivery to embryos through disruption of *GmSWEET10a/b* or *GmSUT1* lowers expression of sucrose metabolism, fatty acid biosynthesis, and TAG assembly genes while inducing storage protein genes, shifting final seed composition toward lower oil and higher protein (Sun et al., 2025). In developing soybean seeds, protein and oil biosynthesis are therefore regulated by distinct but interconnected genetic programs, with storage protein accumulation tied closely to maturation and nitrogen-responsive networks, oil accumulation driven by LEC1/WRI1-centered lipid circuits, and the final balance between the two determined by developmental carbon-nitrogen partitioning.

4 Environmental and Agronomic Factors Affecting Protein and Oil Accumulation Dynamics

4.1 Effects of temperature and climate conditions

Temperature during seed filling strongly shifts the balance between soybean seed protein and oil accumulation, although the response is not strictly linear across environments. Meta-analysis indicates that high temperature generally increases final protein concentration, but this increase does not reflect greater absolute protein synthesis; instead, it arises because other seed fractions, especially oil, are reduced more strongly during stressful filling conditions. Controlled-environment studies are consistent with this pattern, showing that exposure to 35°C during seed fill increased seed protein by 4.0 percentage points and decreased oil by 2.6 percentage points relative to 29°C (Figure 1) (Sun et al., 2025).

The climatic control of composition is also modified by the broader thermal regime and by interactions among multiple weather variables. Large-scale sampling in China showed that crude protein was positively associated with accumulated and mean temperature, whereas crude oil showed the opposite trend overall, but oil increased with mean daily temperature when temperatures remained below 19.7°C, indicating a threshold response rather than a simple monotonic one. Developmental analyses likewise showed that high temperature accelerated maturity,