

and IFS, as well as upstream regulatory factors (Kim et al., 2021; Zhao et al., 2025). Transcriptomic and multi-omics studies have shown that related biosynthetic genes are continuously highly expressed at specific developmental stages in high-isoflavone lines, and isoflavone accumulation is also coordinated with changes in amino acid metabolism, lipid metabolism, and antioxidant defense networks. Therefore, molecular markers linked to or functionally associated with these loci can provide an important basis for the targeted improvement of isoflavone content through MAS.

Fatty acid composition directly determines the nutritional structure, oxidative stability, and processing suitability of soybean oil. High-resolution mapping studies have identified multiple stable QTLs controlling the contents of palmitic acid, stearic acid, oleic acid, linoleic acid, and linolenic acid, including candidate genes such as GmFabG, GmACP, GmFAD8, and fatty acid desaturase-related regions. These QTLs and their tightly linked SNPs can be directly used to screen high-quality oil-type materials, such as those with high oleic acid and low linolenic acid contents (Li et al., 2017; Kumar et al., 2022; Wang et al., 2025).

3.3 Seed nutritional and processing quality

Seed nutritional and processing quality directly determines the application value of soybean in soybean products, plant protein, and feed processing. Its improvement goals have expanded from simply increasing total protein content to optimizing amino acid composition, reducing antinutritional factors, and improving processing suitability. GWAS studies have shown that the contents of essential amino acids such as cysteine, methionine, lysine, and threonine are regulated by multiple QTLs, and some loci are relatively independent of total protein content loci. This means that amino acid balance can be improved through MAS without significantly changing the overall protein level (Lee et al., 2019). In addition, in populations derived from cultivated soybean and wild soybean, some donor alleles that increase protein content have been mapped and shown to function stably with relatively small effects on oil content and agronomic traits, providing traceable targets for backcross introgression and early-generation screening (Huang et al., 2020).

Trypsin inhibitors, lipoxygenases, and certain storage protein variants are important targets for MAS because they affect protein digestibility, beany flavor, oil oxidative stability, and functional properties in food processing, respectively (Kumar et al., 2022; Yao et al., 2022). Molecular markers closely linked to these loci can be used to screen materials with low antinutritional factors, reduced off-flavor, and better processing performance, thereby achieving coordinated improvement of nutritional quality and processing quality. As QTLs related to protein, oil, amino acids, and processing traits are gradually integrated with genomic prediction models, breeders have been able to design ideal genotypes around the comprehensive goals of “high nutritional value-excellent processing characteristics-acceptable agronomic performance” (Sun et al., 2022; Patel et al., 2025).

4 Gene Pyramiding and MAS

4.1 Gene pyramiding strategies

Gene pyramiding refers to a breeding strategy that integrates multiple independent genes or QTLs with favorable effects into the same genetic background, aiming to simultaneously express multiple desirable traits in a single variety, thereby achieving broader, more stable, or more durable improvement effects than individual genes (Dormatey et al., 2020; Haque et al., 2021). In soybean, this strategy is particularly suitable for improving disease and pest resistance, stress tolerance, and certain quality traits, which are determined by multiple functionally complementary major- or moderate-effect loci. Unlike traditional cumulative selection based on phenotypes, MAS can directly track multiple target loci in segregating generations, improving pyramiding efficiency and reducing misselection caused by environmental interference (Das et al., 2017). In disease-resistance breeding, pyramiding multiple R genes or major-effect QTLs, or combining major genes with minor-effect resistance loci, has been shown to enhance resistance spectrum and durability. This principle also applies to complex adaptive traits such as drought, salt, and waterlogging tolerance (Dormatey et al., 2020; Haque et al., 2021).