



## OPEN ACCESS

## EDITED BY

Petr Smýkal,  
Palacký University in Olomouc, Czechia

## REVIEWED BY

Juan Pablo Renzi,  
Instituto Nacional de Tecnología  
Agropecuaria (Argentina), Argentina  
Robert Harrison,  
Murdoch University, Australia

## \*CORRESPONDENCE

Yongwang Sun

✉ sunyongwang@lcu.edu.cn

†These authors have contributed equally to this work

RECEIVED 19 April 2024

ACCEPTED 13 June 2024

PUBLISHED 26 June 2024

## CITATION

Sun Y and Gong Y (2024) Research advances on the hard seededness trait of soybean and the underlying regulatory mechanisms.

*Front. Plant Sci.* 15:1419962.

doi: 10.3389/fpls.2024.1419962

## COPYRIGHT

© 2024 Sun and Gong. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Research advances on the hard seededness trait of soybean and the underlying regulatory mechanisms

Yongwang Sun\*† and Yujie Gong†

School of Agricultural Science and Engineering, Liaocheng University, Liaocheng, China

Soybean is one of the world's most economically significant crops and is widely utilized as an essential source of vegetable protein and edible oil. Cultivated soybean is domesticated from its annual counterpart, wild soybean, which is considered valuable germplasm for soybean breeding. However, wild soybean accessions generally produce seeds with impermeable coats, a trait known as hard seededness (HS), which is beneficial for long-term seed survival but is undesirable for the uniform water absorption and germination of seeds, thus limiting the utilization of wild soybeans in breeding. In addition, moderate HS can isolate the embryo from the surrounding environment and is thus beneficial for long-term seed storage and germplasm preservation. The HS trait is primarily associated with the structure and chemical composition of the seed coat. Moreover, its development is also influenced by various environmental conditions, such as water and temperature. Genetic analysis has revealed that HS of soybean is a complex quantitative trait controlled by multiple genes or minor quantitative trait loci (QTL), with many QTLs and several causal genes currently identified. Investigating the physiological and molecular mechanisms underlying this trait is crucial for soybean breeding, production, and food processing. For this article, the literature was reviewed and condensed to create a well-rounded picture of the current understanding of internal and external factors, QTLs, causal genes, and the regulatory mechanisms related to the HS of soybean, with the aim of providing reference for future research and utilization of this trait.

## KEYWORDS

soybean, hard seededness, seed coat structure, chemical component, QTL, gene

## Introduction

Soybean [*Glycine max* (L.) Merr.] is an important leguminous crop due to its high protein and oil contents. Generally, oil, protein, and carbohydrate account for approximately 20%, 40%, and 35%, respectively, of the dry mass of cultivated soybean seeds, making it an excellent source of vegetable oil and protein for humans and livestock

(Medic et al., 2014; Hudson, 2022). Cultivated soybean was domesticated from the annual wild soybean (*Glycine soja* Sieb. and Zucc.) ~5000 years ago, in the eastern half of north China and disseminated to Europe and America in the 18<sup>th</sup> century (Li et al., 2008; Kim et al., 2012). Wild soybeans are distributed throughout East Asia, including most of China, Korea, Japan, and part of Russia, and have extremely rich genetic resources (Larson et al., 2014). Compared to their wild relatives, cultivated soybeans have undergone significant changes in growth habits, plant architecture, seed morphology, and composition under long-term targeted selection and improvement (Liu et al., 2007; Wen et al., 2009; Yuan et al., 2022). Cultivated soybeans display a bush-type growth habit with a stout primary stem and sparse branches, bearing large seeds that are generally prone to water absorption and germination. In contrast, wild soybeans produce a procumbent or climbing vine with slender, many-branched stems bearing small, coarse seeds that are impermeable to water (Zhou et al., 2010; Liu et al., 2021). During selective domestication and specific breeding improvement, approximately 50% of the genetic diversity and 81% of rare alleles in wild soybeans have been lost (Hyten et al., 2006). At present, most soybean cultivars worldwide face a “genetic bottleneck” (Guo et al., 2010). Many of the soybean cultivars come from a few ancestors (Li et al., 2008; Wymierski and Vello, 2013; Li et al., 2023), and the narrow genetic basis and limited gene sources not only inhibit cultivar improvement, but also increase their vulnerability to abiotic stresses, disease, and pests (Zhou et al., 2015).

Due to not being influenced by human selection, wild soybeans contain abundant genetic resources, which have advantages such as multiple flowers and pods, high protein content, high reproductive coefficients, and tolerance to infertility (Liu et al., 2007; Zeng et al., 2012). In addition, wild and cultivated soybeans exhibit normal pairing of meiotic chromosomes and hybridize easily (Liu et al., 2007; Kim et al., 2012). Thus, wild soybeans are an extraordinarily important gene pool for soybean breeding, which can enrich and broaden the genetic background of cultivated soybeans, and provide beneficial traits and genetic resources (Chen et al., 2017; Kim et al., 2021; Aleem et al., 2022). As reviewed by Zhuang et al. (2022), wild soybeans had been used as genetic resource to improve the biotic and abiotic resistance and yield-related traits of soybean cultivars. However, seeds from wild soybean accessions generally exhibit a hard seededness (HS) trait, which impedes the seeds' water absorption even after long periods of soaking until external factors (e.g. fluctuating temperatures, fire, drying, freezing) render the seed coat permeable to water (Baskin et al., 2000). The HS trait, which has been actively selected against by plant breeders since uniform and rapid emergence is impeded, may appear as a “linkage drag” in hybrid breeding programs, seriously limiting the utilization of wild soybean resources (Sharma et al., 2020). For example, the HS rate of the F<sub>2</sub> and F<sub>3</sub> seeds obtained by crossing the *G. max* line T14R1251–70 and *G. soja* accession LNLT was 89.5% and 96.5%, respectively (Liang et al., 2023). This will undoubtedly increase the difficulty of the subsequent breeding process. By contrast, moderately impermeable seed coats can prevent seed decay and prolong seed lifespan by isolating the embryo from the surrounding environment and is thus beneficial for prolonging the storage time of seeds and germplasm preservation, especially in humid environments

(Hartwig and Potts, 1987; Kulik and Yaklich, 1991; Roy et al., 1995; Tyler, 1997; Moore and Yaklich, 2000). Therefore, breeding programs have been aimed at developing soybean cultivars with seed coats that are reasonably permeable and fairly strong (Kilen and Hartwig, 1978; Moore and Yaklich, 2000; Ma et al., 2004). For example, soybean breeding lines exhibiting HS trait, such as D67–5677, D67–5679, D86–4565, and D87–4647, had been most used in breeding cultivars that are well adapted to the southern U.S.A (Yaklich and Cregan, 1989; Moore and Yaklich, 2000). Understanding the physiological and molecular mechanisms controlling the HS trait of soybean is of great significance for its application. Previous studies have extensively investigated the seed coat structure and components, environmental factors, quantitative trait loci (QTL), causal genes, and regulatory mechanisms associated with HS. This paper aims to provide an overview of the research advances related to the HS trait of soybean and propose future research directions in this field.

## Definition and implications of the HS trait

HS, also known as physical dormancy, seed coat-imposed dormancy, or “stone” or “impermeable” seeds, has been found in ~15 angiosperm families, of which the Leguminosae is the largest (Meyer et al., 2007; Baskin and Baskin, 2014; Paulsen et al., 2014; Wen et al., 2024). In contrast, seeds that are prone to water absorption are termed “soft” or “permeable” seeds. HS is the second most common type of seed dormancy after physiological dormancy and occurs when the seed coat becomes thick or hard, impermeable to surrounding water and air, and keeps the embryo in a viable state for a long time (Foley, 2001; Finch-Savage and Leubner-Metzger, 2006; Paulsen et al., 2013). HS is a biologically beneficial trait that plants have acquired over a long evolutionary process to adapt to environmental changes. For example, seeds with this trait are not easily detected by rodents through their sense of smell, are difficult to digest, and can spread to a larger region as the forager migrates (Nik and Parbery, 1977). After encountering harsh environments such as floods and fires, many of the hard seeds remain viable and can quickly resettle on areas, as long as the environmental conditions are suitable for seed germination (Ferrandis et al., 1999). Hard seeds may benefit the continuation of species in nature by preserving a seed stock for many years so that seeds can germinate at different times to avoid sudden disasters (Li and Huang, 1994; Mohamed-Yasseen et al., 1994; Koornneef et al., 2002). The hard seed coat isolates the embryo from the outside air and water under rainy conditions to maintain the long-term vitality of the seed and avoid pathogen invasion (Dalling et al., 2011). For example, some legume seeds have remained viable for more than 100 years (Rolston, 1978).

Like other plants, hard seeds from wild soybean accessions have higher vitality and can survive for at least five to seven years in the soil seed bank, which is beneficial for adaptation to unstable and unpredictable environmental conditions to ensure population continuity (Li and Huang, 1994; Heatherly et al., 1995). The longevity of wild soybeans can be attributed to their combined physical and physiological dormancy (Sun et al., 2015; Wang et al., 2018). In contrast,

most soybean landraces and cultivars can imbibe water in a short time without scarification of the seed coat, although considerable variation exists in their degree of seed coat permeability (Chachalis and Smith, 2000; Sakamoto et al., 2004; Kebede et al., 2014). For example, hard seed ratio varied by an average of 15% for eleven strains in Mississippi (Hartwig and Potts, 1987). Mullin and Xu (2001) reported that three of the six tested soybean cultivars from U.S.A., Canada and Japan produce hard seeds with varying degrees (e.g. Bobcat, 3.1%; Harosoy, 5.6%; OX 951, 72.4%), while the others all produce non-hard seeds. Guo et al. (2002) reported that five of the seven test soybean cultivars from northeast China produce a relatively small proportion of hard seed (HS ratio ranging from 0.5% - 3.0%). Moreover, adverse climatic conditions such as drought and heat during the growing season, especially as the seed matures, may increase the ratio of hard seeds in soybean production (Moore and Yaklich, 2000; Mullin and Xu, 2001). Currently, HS is often a disadvantageous trait for cultivated soybeans and seriously impedes seed production in agriculture (Meyer et al., 2007). Sowing of hard soybean seeds leads to a dramatic decrease in field emergence rate and uniformity, resulting in different seedling density and mass weed growth, which seriously decreases the sowing quality and ultimately reduces the soybean yield (Potts et al., 1978; Tyler, 1997). Under the continuous cropping system, if hard seeds germinate the following year, they will mix with subsequent soybean seeds (sometimes from different cultivars), and the harvest quality will be decreased (Keith and Delouche, 1999; Meyer et al., 2007; Abbo et al., 2008). In the food industry, rapid and uniform hydration of soybean seeds is necessary for subsequent seed processing. Hard seeds will not only affect the sensory quality of whole-soy foods, including soybean sprouts, natto, and baked beans, but also reduce the milling quality of crack-soy foods such as soy milk, soy sauce, soy paste, miso, and tofu (Zhang et al., 2008; Smýkal et al., 2014). To ensure uniform water absorption and germination of soybean seeds with the HS trait, multiple methods, such as mechanical scratching, temperature upheaval, or soaking seeds with strong acid or alkaline solutions, have to be employed to break their seed coat (Argel and Paton, 1999; Moore and Yaklich, 2000; Shao et al., 2007; Sun et al., 2014). Undoubtedly, these steps increase the processing costs and may also decrease the seed vitality.

In addition to the adverse effects on agricultural production and food processing, the HS trait has also shown valuable utilization in certain agricultural practices. For example, if adverse weather (e.g. humid conditions or rainy season) occurs before seed maturity or harvest, the hard seed coat can hinder the diffusion of water and air to the embryo to prevent seed spoilage. Different with wild soybean, cultivated soybean seeds generally lack physiological dormancy and are prone to deterioration in seed viability under prolonged storage (Qutob et al., 2008; Ramakrishna et al., 2018; Wang et al., 2018). Thus, a moderate or rational level of HS can delay the embryo's water absorption from the surrounding environment, resist pathogen invasion, and help the seed maintain strong vitality and good quality for a longer period (Potts et al., 1978; Keim et al., 1990; Kulik and Yaklich, 1991; Roy et al., 1995; Wang, 1999). In addition, HS may be associated with high calcium content in the seed coat and can therefore potentially enhance the nutritional value of soy-based foods (Saio et al., 1973; Saio, 1976; Chen et al., 2001; Zhang et al., 2009; Sun et al., 2015). Therefore, a deep understanding of the

physiological and molecular basis of the HS trait is necessary for the rational application of wild soybean resources and the HS trait in soybean cultivar improvement, soybean production, and soy food processing practices.

## Internal and external factors affecting the HS trait

HS in soybean is genetically and environmentally determined. In terms of soybean seed characteristics, HS is mainly related to the structure and chemical composition of the seed coat (Ma et al., 2004; Meyer et al., 2007; Qutob et al., 2008), although several studies have emphasized the role of the hilum or micropyle in water absorption (Hyde, 1954; McDonald et al., 1988; Pietrzak et al., 2002; Xu et al., 2009). The seed coat, the outer covering of the seed, is largely derived from the integuments of the maternal tissue after fertilization, consisting of an embryo and nutritive tissue (Boesewinkel and Bouman, 1995). It can not only transfer nutrients from the mother plant to the embryo during seed development but also protects the embryo from mechanical damage or pathogen invasion (Souza and Marcos-Filho, 2001; Brooker et al., 2007). Meyer et al. (2007) suggested that the imbibition process by soybean seeds occurs in two distinct phases, the first dominated by hydration of the seed coat and the second by hydration of the cotyledons, which is rate-limited by the former. After removal of the seed coats, no significant difference was found in the water absorption rates between permeable and impermeable seeds. The water permeability rate of the hard seed coat was found to be five times lower than that of the permeable seed, indicating that seed coat permeability is the main factor leading to the HS trait. Moreover, permeable seeds typically imbibed water initially through their dorsal sides, opposite the hilum, forming wrinkles in their seed coats and delivering water to the underlying cotyledons (Meyer et al., 2007). The soybean seed coats are covered by a cuticle, which is subtended by closely packed palisade macrosclereids (also known as Malpighian layer cells), thick-walled osteosclereids (also known as hourglass cells), crushed parenchyma, and an intact aleurone layer (Yaklich et al., 1992; Shao et al., 2007). The cuticle is a structural component of the seed coat that provides a potential barrier to water movement (Heredia, 2003). Using scanning electron microscopy (SEM), Ma et al. (2004) observed the surface structure of seed coats from six soybean cultivars exhibiting different water imbibition rates. They found that the cuticle of permeable seeds has many microscopic cracks that typically penetrate the outer periclinal walls of the palisade layer, which occur more frequently on the dorsal side than in other regions. In contrast, the cuticle of hard seeds (*G. max* cv. OX951, from U.S.A.) is relatively intact on the whole of the seed coat surface. Moreover, isolated pieces of cuticle from hard seeds were found to be stronger and less prone to breaking when handled than those isolated from permeable seeds. After immersion in water, the typical water absorption phenomenon of seed coat wrinkles are first observed on the dorsal side, and then the wrinkles extend to the lateral side. In contrast, a fluorescent tracer

dye assay showed that the hilum, micropyle, and raphe were not sites of initial water entry. Therefore, the authors suggested that the integrity of the cuticle and its underlying palisade layer at the dorsal site is the key factor that determines the permeability of a soybean seed coat (Ma et al., 2004). A time-course observation of the seed coat from the developing seeds of different soybean cultivars showed that at the L2b stage, when seed expansion was maximal, cracks in the cuticle start to appear on the dorsal region of permeable seeds, but not impermeable seeds, regardless of their genotype (Ranathunge et al., 2010). These results further confirm that it is the outermost cuticle, especially when located on the dorsal side of seeds, that controls water flow into soybean seeds during imbibition. In contrast, the determining factors of seed coat permeability in wild soybeans may differ from those in cultivated soybean seeds exhibiting HS trait. The micropyle of wild soybean seeds cannot be observed under SEM due to them being covered by the cuticle. However, after treatment with concentrated sulfuric acid, which leads to the shedding of the cuticle, the micropyle of wild soybean seeds is exposed and the seeds become permeable (Xu et al., 2009). Thus, Xu et al. (2009) believed that the micropyle may be the main channel for water uptake. Additional experiments, such as the fluorescent tracer dye assay conducted by Ma et al. (2004), were needed to determine the initial site of water entry in wild soybeans.

In addition to the seed coat structures, the presence of certain chemicals on the seed coat are also related to the HS trait. Mullin and Xu (2001) determined the chemical components of the seed coats of six soybean cultivars, and found that there was no correlation between the concentration of any of the cations and the occurrence of HS, although many other studies have shown that there is a close relationship between the calcium content of the seed coat and the HS trait of cultivated soybean (Saio et al., 1973; Saio, 1976; Chen et al., 2001; Zhang et al., 2009; Sun et al., 2015). Interestingly, low pectin content and high hemicelluloses, particularly xylans, can reduce hydrophilicity and may lead to the occurrence of HS (Mullin and Xu, 2001). Shao et al. (2007) found that the outermost cuticle of hard seed coats contained a high amount of hydroxylated fatty acids relative to that of permeable seeds. Moreover, boiling hard wild soybean seeds in 1 M NaOH for 5 min released the  $\omega$ -hydroxy fatty acid component of the cuticle and created holes in the seed coat surface, causing the seeds to become permeable. Three major phenolics, namely epicatechin, cyanidin-3-O-glucoside, and delphinidin-3-O-glucoside, were isolated from wild soybean seed coats, and the changes in epicatechin content were significantly positively correlated with the HS trait under different water conditions during seed development and under different gas conditions during seed storage (Zhou et al., 2010). Similarly, Vijayan et al. (2023) measured the phenols, tannins, proteins, trace elements, and metabolites in the seed coats of two soybean cultivars exhibiting different levels of water permeability. They found that phenols are the main compounds responsible for seed coat impermeability, which may provide strength to the seed coat.

Previous studies have also shown that the HS trait of soybean may also be related to seed color and size and is influenced by

multiple environmental factors. Inheritance studies have found that there may be an association between pigmentation loci and HS in soybean, where dark-colored seeds determined by genetic factors generally exhibit lower imbibition rates than light-colored seeds (Kuo, 1989; Sakamoto et al., 2004; Watanabe et al., 2004; Liu et al., 2007). However, several other studies have shown that seed coat pigmentation is not closely linked with HS, since many soybean lines without coloring possess an impermeable seed coat (Keim et al., 1990; Chachalis and Smith, 2000; Ma et al., 2004). Qutob et al. (2008) suggested a possible pleiotropic effect between pigmentation loci and permeability through their activity toward polyphenolic substrates that become impregnated in the cell wall. Small seeds, regardless of genotype, are found to be positively associated with HS of soybean (Calero et al., 1981; Nooden et al., 1985; Yaklich et al., 1986; Hill et al., 1986a; Ragus, 1987; Taira, 1990; Moore, 1991; Huang et al., 1999). Adverse growing conditions such as drought, low moisture, and high temperature, especially at the seeds' late developmental stage, typically increase the occurrence of hard seeds in soybean; therefore, the hard seed ratio in cultivated soybean often vary between different years and geographic locations (Baciu-Miclaus, 1970; Hill et al., 1986b; Vieira et al., 1992; Argel and Paton, 1999; Egli et al., 2005; Kebede et al., 2014; Jaganathan and Harrison, 2024). As observed by Ranathunge et al. (2010), the cuticular cracks that are critical for seed permeability were induced by the internal hydrostatic forces of cotyledons when seed expansion was maximal. The reason why seeds with small sizes and those that encounter adverse climates are typically impermeable is probably because their cotyledons are unable to continue to expand rapidly later in development and are not prone to result in cuticular cracking.

## QTL mapping of HS trait in soybean

The localization and cloning of genes or QTLs for important traits are of great significance in crop breeding. In the past few decades, researchers have used many  $F_2$  populations and/or recombinant inbred lines (RILs) to perform extensive linkage analysis of QTLs associated with the HS trait of soybean. In general, researchers have found that the HS trait is a dominant trait in soybean, which is controlled by multiple genes or minor QTLs, although several papers have reported the monogenetic control of this trait (Kebede et al., 2014; Sun et al., 2015). Initially, Kilen and Hartwig (1978) studied the inheritance of HS in soybeans by crossing two soybean breeding lines, 'Tracy' and 'D67-5679' (producing 88% hard seed) from U.S.A., with different levels of seed coat permeability, and they speculated that as few as three major genes may control the variation in soybean HS. Using a  $F_2$  segregating population derived from a cross between *G. max* and *G. soja*, Keim et al. (1990) identified five restriction fragment length polymorphism markers, located on chromosomes 2, 3, 8, and 19, that were associated with variations in the HS trait. These markers and their epistatic interactions accounted for a total of 71% of variation in HS. Using a similar strategy, Sakamoto et al. (2004) found that at least two major QTLs and one minor QTL were

involved in controlling the HS trait. Among them, a QTL near Satt459, a simple sequence repeat (SSR) marker on chromosome 2, could explain 23.8% and 38.5% of the phenotypic variations in the F<sub>2</sub> and F<sub>3</sub> generations, respectively. Using an F<sub>2</sub> population derived from two soybean cultivars (non-HS cv. JS 71-05 and HS cv. Brisa soya-1 from India) exhibiting different levels of seed coat permeability, Singh et al. (2008) found that four SSR markers, located on chromosomes 6, 8, 12, and 15, were significantly correlated with seed coat permeability, explaining 3.9%–4.5% of the total phenotypic variation. Among these markers, Satt281 also showed significant association with electrolyte leaching and explained 5.6% of the total variation for this trait, which was negatively associated with seed vigor during germination. Based on two F<sub>2</sub> populations derived from a cross and reciprocal cross made between two *G. max* accessions (non-HS cv. PI 587982A and HS cv. PI 594619 from U.S.A) exhibiting different levels of seed coat permeability, Kebede et al. (2014) found that the main locus controlling soybean seed coat hardness in *G. max* accession PI594619 is located in a 0.18-Mb interval between the markers Satt202 and Satt459 on chromosome 2.

Compared to the F<sub>2</sub> segregating population, the advantage of using a RIL population is that replicated experiments at different locations and/or times can be carried out and enable researchers to evaluate the environmental effects. Based on a RIL population derived from a cross between *G. max* and semi-wild *G. max* lines, Watanabe et al. (2004) identified three QTLs (*RAS1-3*) on chromosomes 2, 6, and 20, which accounted for 10.6%, 26.0%, and 6.1% of the phenotypic variance of HS, respectively. Using a RIL population developed from a cross between *G. max* and *G. soja*, Liu et al. (2007) found two HS QTLs on chromosomes 2 and 6, which were consistently detected over 2 years, accounting for more than 42% and 13% of the phenotypic variations, respectively. Based on a RIL population developed from a cross between a permeable soybean cultivar and an impermeable soybean landrace, Ai et al. (2018) identified three QTLs for HS on chromosomes 2, 6, and 14, respectively, with a genetic contribution rate of 5.54%–12.94%. Additionally, four pairs of epistatic interaction QTLs were detected on chromosomes 2, 6, 9, 12, and 14, respectively, which explained 2.53%–3.47% of the phenotypic variation. Their results indicated that the epistatic interaction effect may play an important role in the genetic basis of the HS trait of soybean. Chen et al. (2019) constructed F<sub>2</sub> and RIL populations by crossing a soybean cultivar with a wild soybean accession. Based on bulked segregant analysis and traditional linkage mapping, three HS QTLs were detected on chromosomes 2, 3, and 6, respectively, which explained 4.9%–23.3% of the phenotypic variation. Chandra et al. (2020) constructed 204 RILs by crossing two wild soybean accessions with a permeable soybean cultivar, and seven HS QTLs were mapped on chromosomes 2, 5, 12, 13, and 16, respectively, explaining 5.96%–39.67% of the phenotypic variations. Two main and stable QTLs distributed in tandem on chromosome 2 were consistent over the 3 years of testing, and they jointly explained 43.09%–62.92% of the phenotypic variations of the HS trait. The QTLs mapped and the linked markers identified in these studies are summarized in Table 1 and Figure 1, which would be useful in

breeding soybean cultivars with rational seed coat permeability by suitable genetic modifications.

## Genes related to HS in soybean

Although researchers have discovered a large number of QTLs related to soybean HS, it is still difficult to apply these QTLs to breeding due to the lack of accuracy and stability. Cloning causal genes underlying the HS trait of soybean and analyzing their regulatory mechanisms is of great significance for the utilization of HS trait. At present, three genes have been confirmed to regulate the formation of the HS trait in soybean seeds. Many studies have reported the presence of a common QTL on an overlapping region of soybean chromosome 2 (Keim et al., 1990; Watanabe et al., 2004; Liu et al., 2007; Zhang et al., 2008; Kebede et al., 2014). To understand the molecular basis of HS, Sun et al. (2015) developed two F<sub>2</sub> populations by crossing the water permeable soybean cultivar Williams 82 and two *G. soja* accessions, and they delimited the QTL to a 22-Kb region containing two genes. Among them, a single nucleotide polymorphism (SNP; C to T at the 1,154<sup>th</sup> site of its coding sequence) in *Glyma02g43700* (*Glyma.02g269500* in Wm82.v4), which could effectively distinguish between the parental genotypes and eight additional *G. soja* accessions, was found to be responsible for the occurrence of HS. *Glyma02g43700* encodes a calcineurin-like metallophosphoesterase transmembrane protein that is primarily expressed in the Malpighian layer of the seed coat, and the amino acid switch (Thr to Met at the 385<sup>th</sup> site of its protein sequence) resulting from the gene mutation was predicted to be located outside the membranes and to affect the protein structure, and thus increased the seed coat permeability (Sun et al., 2015). Jang et al. (2015) developed a near-isogenic line (NIL) of a permeable cultivar containing the HS allele qHS1 from wild soybean. Phenotypic analysis revealed that the qHS1 gene can increase the amount of  $\beta$ -1,4-glucans in the outer layer of palisade cells of the seed coat on the dorsal side of seeds, known to be a point of entrance for water (Ma et al., 2004; Meyer et al., 2007). Fine mapping and soybean transformation confirmed that an endo-1,4- $\beta$ -glucanase gene, *Glyma02g43680* (*Glyma.02g269400* in Wm82.v4), was the causal HS gene. A point mutation (G to T at the 863<sup>rd</sup> site of its coding sequence) introduced an amino acid substitution (Ser to Iso at the 288<sup>th</sup> site of its protein sequence) in a substrate-binding cleft of the enzyme, possibly reducing or eliminating its affinity for substrates in permeable cultivars from various regions, while the allele from the HS genotype was found to promote the accumulation of  $\beta$ -1,4-glucan derivatives that reinforce the HS trait (Jang et al., 2015). The above two genes are both located within the previously reported chromosome 2 interval and are adjacent in position (Sakamoto et al., 2004; Liu et al., 2007; Kebede et al., 2014; Chen et al., 2019; Chandra et al., 2020) and may be related to the strength of the palisade layer of the seed coat. Wang et al. (2022) found that a polygalacturonase gene *PG031* (*Glyma.06G207300*) was specifically expressed in the flowers and reproductive tissues of soybean plants, and that it was strongly

TABLE 1 List of QTLs related to the HS trait of soybean.

Population information <sup>#</sup>	Chromosome	QTL	Marker(s)	Location	LOD	PVE (%)	Reference
Type: F <sub>2</sub> P <sub>1</sub> : Tokei 780, <i>G. max</i> , Japan P <sub>2</sub> : <b>B01167</b> , <i>G. soja</i> , Japan	6	NG	Satt316	48016322	NG	24.8	Sakamoto et al., 2004
	6	NG	Satt489	23848501	NG	8.6	
	2	NG	Satt459	45311085	NG	23.8	
	18	NG	Satt288	51127425	NG	9.1	
Type: RIL P <sub>1</sub> : Misuzudaizu, <i>G. max</i> , Japan P <sub>2</sub> : <b>Moshidou Gong 503</b> , <i>G. max</i> , Japan	2	RAS2	Sat289b	47042650	6.3	10.6	Watanabe et al., 2004
	6	RAS1	Satt489-Satt100	23848501–31490622	14	26	
	20	RAS3	Satt367	2615587	3.0	6.1	
Type: RIL P <sub>1</sub> : Tokei 780, <i>G. max</i> , Japan P <sub>2</sub> : <b>Hidaka 4</b> , <i>G. soja</i> , Japan	6	qHS-C2	Sat_076-Satt307	16007694–46820673	5.9	18.5	Liu et al., 2007
	2	qHS-D1b	Satt459-Staga002	45311085–46281147	12.5	42.5	
Type: F <sub>2</sub> P <sub>1</sub> : JS 71-05, <i>G. max</i> , India P <sub>2</sub> : <b>Brisa soya-1</b> , <i>G. max</i> , India	12	NG	Satt434	38428476	NG	3.9	Singh et al., 2008
	8	NG	Satt538	47395378	NG	4.5	
	6	NG	Satt281	6529270	NG	4.1	
	15	NG	Satt598	13653981	NG	4.2	
Type: F <sub>2</sub> P <sub>1</sub> : PI 587982A, <i>G. max</i> , U.S.A. P <sub>2</sub> : <b>PI 594619</b> , <i>G. max</i> , U.S.A.	2	Isc	Satt459-Sat_202	45311085–45497649	88.1	65.6	Kebede et al., 2014
Type: RIL P <sub>1</sub> : Jidou 12, <i>G. max</i> , China P <sub>2</sub> : <b>ZDD03651</b> , <i>G. max</i> , China	2	qHS-2-1	Sat_069-Sat_183	43278769–44317044	2.51	5.54	Ai et al., 2018
	6	qHS-6-1	Sat_402-Satt460	16418553–44049891	6.64	12.94	
	9	qHS-9-1	Sat_399-Satt273	9095532–38799271	7.24	3.36	
	12	qHS-12-1	12_1-Satt353	18935–1687387	5.87	3.13	
	14	qHS-14-1	Satt577-Sat_287	675214–5287913	3.34	8.25	
Type: F <sub>2</sub> and RIL P <sub>1</sub> : ZH 39, <i>G. max</i> , China P <sub>2</sub> : <b>NY27-38</b> , <i>G. soja</i> , China	2	NG	Satt274-Sat_198	45267040–45542863	13.3	17.2	Chen et al., 2019
	6	NG	6_0993-6_1068	18902966–20942203	13	17.8	
	3	NG	Sat_266-Satt_236	34016929–35595364	2.7	4.9	
	6	NG	Sat_402-Satt557	16418553–20218893	11.5	20.4	
Type: F <sub>2</sub> P <sub>1</sub> : JS 335, <i>G. max</i> , India P <sub>2</sub> : <b>PI 424079 or PI 136620</b> , <i>G. soja</i> , India Type: RIL P <sub>1</sub> : DS 9712, <i>G. max</i> , India P <sub>2</sub> : <b>DC 2008-1</b> , <i>G. soja</i> , India	2	qScI-h 2-1	Satt703-Satt274	42580676–45267040	6.32	16.49	Chandra et al., 2020
	2	qScI-h 2-2	Satt274-Sat_202	45267040–45497649	11.83	26.66	
	5	qScI-h 5-1	Satt619-Satt545	35971621–36463025	3.42	5.96	
	12	qScI-h 12	Satt181-Satt434	36365787–38428476	3.85	6.22	
	13	qScI-h 13	Satt325-Satt252	13272992–16454986	3.01	6.71	
	16	qScI-h 16	Sct046-Satt686	2919448–13053666	3.72	6.74	

<sup>#</sup>Parent names in normal and bold fonts produce seeds with non-HS and HS trait, respectively. QTL, Quantitative trait loci; NG, the information was not given in the articles; LOD, logarithm of the odds; PVE, percentage of phenotypic variation.

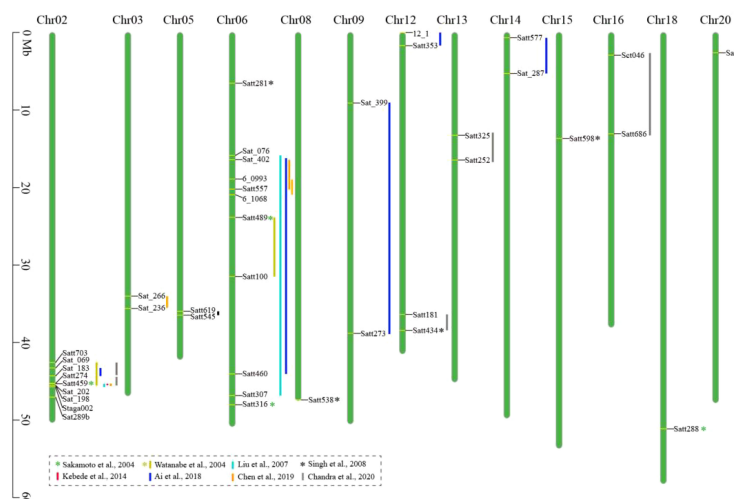


FIGURE 1

Chromosomal locations of the hard seededness (HS) quantitative trait loci (QTLs) of soybean.

induced when the seed coat absorbs water. *PG031* underlies a QTL on chromosome 6 that can explain ~15% of phenotypic variation in seed coat permeability (Liu et al., 2007), and it has three haplotypes (289H, 289Y, Hap3) in the soybean population. Overexpression of the impermeable allele *PG031*<sup>289H</sup> in the permeable cultivar reduced the water permeability of transgenic seeds by reducing the level of pectins and decreasing intracellular spaces of the osteoscleroid layer and parenchyma of the seed coat to decline water accessing the seed, and the 100-seed weight was also decreased. A genetic diversity analysis in a diverse soybean panel suggested that the permeable allele *PG031*<sup>289Y</sup>, associated with high water permeability and high seed weight, is experiencing ongoing artificial selection (Wang et al., 2022). In addition, two studies have predicted that several genes may be associated with HS of soybean based on comparative omics analysis. Whole-genome resequencing was conducted to identify the SNPs and InDels between genomes of cultivated and wild soybean in relation to their HS trait. Seven genes carrying nonsynonymous variants were found to play a probable role in influencing HS. Two of them (genes encoding type-I inositol polyphosphate-5-phosphatase-1 and E3 ubiquitin ligase) were found to segregate among the RILs in coherence with their water permeability scores and showed a preliminary association with the desirable water permeability characteristics (Ramakrishna et al., 2018). Recently, Wang et al. (2023) conducted transcriptomics and proteomics analysis of a hard-seeded chromosome segment substitution line R75 and its recurrent parent SN14, and multiple genes were found to be associated with HS. Five of them were selected as candidates based on their expression patterns, functional annotations, SNPs, and protein structural changes. These genes may play important roles in the formation of the cell wall and seed coat, and showed significantly higher expression in seeds of R75 than in those of SN14. Competitive allele-specific PCR (KASP) markers were developed on the basis of their nonsynonymous SNP and showed high selection efficiency in distinguishing hard and nonhard soybean lines. The precise functions and regulatory

mechanisms of these genes in regulating the HS trait require further research.

## Physical dormancy genes in other genera

In addition to those in soybean, genes in other plants have been found to regulate the level of seed coat permeability. In *Arabidopsis*, for example, mutants with altered proanthocyanidin production in the seed coat, such as *ban* (Albert et al., 1997) and *transparent testa 1, 10, and 13* (Debeaujon et al., 2000), and those defective in the secondary cell walls of the outer layers, such as *ap2* (Western et al., 2001), produce seeds that are more permeable than those of wild-type plants. Mutations of the laccase gene, *ATLAC15*, resulted in a reduction in lignin content and elevated permeability of *Arabidopsis* seed coats (Liang et al., 2006). An acyl-CoA:glycerol-3-phosphate acyltransferase gene, *ATGPAT5*, is required for the synthesis of suberin in the seed coat and root of *Arabidopsis*, and mutations in this gene result in a reduction in the very long chain dicarboxylic acids and  $\omega$ -hydroxy fatty acids typical of suberin and a steep increase in seed coat permeability (Beisson et al., 2007). *Medicago truncatula* is a model plant of the legume family due to its small genome, short growth period, and high genetic transformation efficiency, and it generally produces hard seeds. By screening of Tnt1 retrotransposon-tagged *Medicago truncatula* lines, Chai et al. (2016) identified mutants that produce permeable seeds. A class II KNOTTED-like homeobox (KNOXII) gene, *KNOX4*, was found to be responsible for the loss of the HS trait in the seeds of the mutants. Mutation of *KNOX4* altered seed coat cuticle composition, and the *knox4* mutant developed a structurally anomalous cuticle. Subsequent experiments showed that *CYP86A*, a gene associated with cutin biosynthesis, is a downstream target of *KNOX4*. Furthermore, the expression level of a seed coat-expressed  $\beta$ -ketoacyl-CoA synthase gene, *KCS12*, was found to be regulated by

KNOX4. The concentration of C24:0 lipid polyester monomers are significantly decreased in *kcs12* mutant seeds, which exhibit irregular seed coats and lose the HS trait (Chai et al., 2021). These findings reveal a molecular mechanism by which KNOX4 and KCS12 control development of the seed coat and the HS trait. Common bean (*Phaseolus vulgaris*) is an important legume crop for human consumption, and some cultivars still retain a high level of the HS trait. To investigate the molecular basis of the rate of seed water absorption, an RIL population segregating for HS was constructed, and the *pectin acetyltransferase 8* gene was found to be critical for the HS trait of common bean (Soltani et al., 2021). The nonfunction allele of *pectin acetyltransferase 8* may be responsible for the development of permeable seeds, and was formed under strong selection pressure through domestication. In *Vigna stipulacea* Kuntze, another legume crop that inhabits mainly in South Asia, researchers have found that a cellulose synthase, a catalytic subunit gene *VsCESA7*, and a phospholipid sterol acyltransferase gene *VsPSAT1* may be responsible for the development of HS in their wild species (Takahashi et al., 2019; 2023). These studies show that the biosynthesis of some key chemical components is critical for seed coat permeability, and provide insights for investigation into soybean's HS trait.

## Summary and future prospects

HS is a type of seed dormancy found in many angiosperms, which isolates the embryo from the external environment through its hard or impermeable seed coat, thus serving as an important mechanism for ensuring the survival and perpetuation of the species. This trait is particularly evident in leguminous plants, for example, wild soybeans, producing hard seeds. During domestication and breeding processes, most cultivated soybeans have lost the HS trait, which is beneficial for soybean production and processing. However, several cultivated soybean cultivars (e.g. HS cultivars in U.S.A., Japan and India, which has been used for QTL mapping of the HS trait, see Table 1) still exhibit the HS trait, and the expression of this trait is often influenced by environmental factors. Unveiling the physiological and molecular regulatory mechanisms governing the hard seed trait in soybeans is of significant importance for the utilization of wild soybeans in breeding, preservation of germplasm resources, and soybean production and processing. Previous studies on the HS trait in soybeans have made significant progress, and this paper provides a summary of this knowledge. The HS break of soybean seeds is mainly associated with microscopic cracks in the seed coat's cuticle layer. The composition and content of fatty acids, phenols, and xylans may be related to the strength of the cuticle layer, thus participating in the formation of the HS trait. To date, nearly 30 QTLs related to the HS trait have been discovered on 13 chromosomes of soybeans, with significant differences in their contributions to phenotypic variation. These QTLs are very useful in molecular marker-assisted selection breeding in soybean and explaining the molecular mechanism underlying the HS trait.

Among them, QTLs on chromosomes 2 and 6 have been repeatedly reported by different researchers, who have identified three causal genes responsible for HS based on this information. Through comparative omics analyses, some genes are also believed to be associated with the HS trait. In this article, we summarized the reported HS-related genes and their regulatory mechanisms in Arabidopsis and three legumes (*Medicago truncatula*, common bean, and *Vigna stipulacea* Kuntze), which mainly regulate HS by modulating the composition and content of the seed coat.

The above information provides insights into the mechanisms and utilization of the HS trait in soybeans, while there are still many unresolved issues in this research field. First, why do adverse or harsh environmental conditions induce the HS ratio in some soybean cultivars? Are there other regulatory mechanisms besides limiting the expansion of mature soybean cotyledons? For instance, do environmental factors affect the expression levels of certain genes in the seed coat, thus directly influencing seed coat composition that contributes to HS? Second, besides substances already reported by previous studies, what other key substances determine the strength and permeability of soybean seed coats? How is the synthesis of these substances regulated, and how do they affect seed coat structure? Studying HS and non-HS soybean seed coat caused by genetic or environmental factors through multi-omics techniques (e.g. transcriptome, proteomics, metabolomics) and analyze its regulatory mechanism through biochemical and molecular biology experiments may provide some clues to the aforementioned questions. Third, multiple HS QTLs have been discovered so far. Apart from chromosomes 2 and 6, which other genes on other chromosomes are involved in the formation of the HS trait in soybeans? Given that the HS trait is controlled by multiple genes, constructing genetically stable RIL or NIL populations by single seed descent or backcrossing method, respectively, will facilitate the cloning of other HS genes and the elucidation of regulatory mechanisms. Finally, how can currently cloned HS genes be utilized to breed soybean cultivars with moderately or fairly strong seed coats? Application of these cultivars can not only maintaining soybean production and food processing but also extending the storage time of soybean seeds, especially in humid regions. Modifying the expression levels and biochemical activities of HS regulatory genes/proteins through genetic engineering and gene editing techniques may be a preferable approach.

## Author contributions

YS: Funding acquisition, Investigation, Writing – original draft, Writing – review & editing. YG: Writing – original draft.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was financially supported by National Natural Science Foundation of



China (32101788), the Industrial Promotion Project of Shandong Agricultural Science and Technology Park (2019YQ035), the Open Project of Liaocheng University Landscape Architecture Discipline (319462212), the Higher Educational Program of Shandong Province for Introduction and Cultivation of Young Innovative Talents (S202210447038; CXCY2023264), and the Doctoral Research Start-up Foundation of Liaocheng University (318052021).

## Acknowledgments

We appreciate the linguistic assistance provided by ZYEdit during the preparation of this manuscript.

## References

- Abbo, S., Zezak, I., Schwartz, E., Lev-Yadun, S., and Gopher, A. (2008). Experimental harvesting of wild peas in Israel: implications for the origins of near east farming. *J. Archaeol. Sci.* 35, 922–929. doi: 10.1016/j.jas.2007.06.016
- Ai, L. J., Chen, Q., Yang, C. Y., Yan, L., Wang, F. M., Ge, R. C., et al. (2018). Mapping main-effect and epistatic QTL for hard seediness in soybean. *AAS* 44, 825–828. doi: 10.3724/SP.J.1006.2018.00852
- Albert, S., Delsene, M., and Devic, M. (1997). BANYLUS, a novel negative regulator of flavonoid biosynthesis in the Arabidopsis seed coat. *Plant J.* 11, 289–299. doi: 10.1046/j.1365-313X.1997.11020289.x
- Alem, M., Alem, S., Sharif, I., Alem, M., Shahzad, R., and Khan, M. I. (2022). Whole-genome identification of APX and CAT gene families in cultivated and wild soybeans and their regulatory function in plant development and stress response. *Antioxidants (Basel)* 11, 1626–1651. doi: 10.3390/antiox11081626
- Argel, P., and Paton, C. (1999). “Overcoming legume hardseededness,” in *Forage Seed Production: Tropical and Subtropical Species*, vol. 2. Eds. D. S. Loch and J. E. Ferguson (CAB International, Wallingford), 247–265.
- Baciu-Miclaus, D. (1970). Contribution to the study of hard seed and coat structure properties of soybean. *Proc. Int. Seed Test. Assoc.* 35, 599–617.
- Baskin, J. M., Baskin, C. C., and Li, X. J. (2000). Taxonomy, anatomy and evolution of physical dormancy in seeds. *Plant Species Biol.* 15, 139–152. doi: 10.1046/j.1442-1984.2000.00034.x
- Baskin, C. C., and Baskin, J. M. (2014). *Seeds: Ecology, Biogeography and Evolution of Dormancy and Germination*. 2nd Edn (New York: Academic Press).
- Beisson, F., Li, Y., Bonaventure, G., Pollard, M., and Ohlrogge, J. B. (2007). The acyltransferase GPAT5 is required for the synthesis of suberin in seed coat and root of Arabidopsis. *Plant Cell* 19, 351–368. doi: 10.1105/tpc.106.048033
- Boesewinkel, F. D., and Bouman, F. (1995). “The seed: structure and function,” in *Seed Development and Germination*, eds. J. Kigel and G. Galili (New York: Marcel Dekker), 1–24. doi: 10.1201/9780203740071
- Brooker, N. L., Lagalle, C. D., Zlatanic, A., Javni, I., and Petrovic, Z. (2007). Soy polyol formulations as novel seed treatments for the management of soil-borne diseases of soybean. *Commun. Agric. Appl. Biol. Sci.* 72, 35–43.
- Calero, E., West, S. H., and Hinson, K. (1981). Water absorption of soybean seeds and associated causal factors. *Crop Sci.* 21, 926–933. doi: 10.2135/cropsci1981.0011183X002100060030x
- Chachalis, D., and Smith, M. L. (2000). Imbibition behavior of soybean (*Glycine max* (L.) Merrill) accessions with different testa characteristics. *Seed Sci. Technol.* 28, 321–331.
- Chai, M., Queralt Castillo, I., Sonntag, A., Wang, S., Zhao, Z., Liu, W., et al. (2021). A seed coat-specific  $\beta$ -ketoacyl-CoA synthase, KCS12, is critical for preserving seed physical dormancy. *Plant Physiol.* 186, 1606–1615. doi: 10.1093/plphys/kiab152
- Chai, M., Zhou, C., Molina, I., Fu, C., Nakashima, J., Li, G., et al. (2016). A class II KNOX gene, *KNOX4*, controls seed physical dormancy. *P Natl. Acad. Sci. U.S.A.* 113, 6997–7002. doi: 10.1073/pnas.1601256113
- Chandra, S., Taak, Y., Rathod, D. Y., Yadav, R. R., Poonia, S., Sreenivasa, V., et al. (2020). Genetics and mapping of seed coat impermeability in soybean using inter-specific populations. *Physiol. Mol. Biol. Plants* 26, 2291–2299. doi: 10.1007/s12298-020-00906-y
- Chen, P., Huhn, M. R., Buss, G. R., Gunduz, I., and Wilson, J. H. (2001). Interrelationship among seed quality traits of specialty soybeans for natto and tofu. *ASA-CSSA-SSSA Int. Annu. Meet.*, (Charlotte, NC).
- Chen, J., Liu, X., Yu, L., Lu, Y., Zhang, S., Zhang, H., et al. (2019). QTL mapping of hard seediness in wild soybean using BSA method. *Sci. Agric. Sin.* 52, 2208–2219. (in Chinese).
- Chen, X. F., Ning, K., Xu, H. L., Yu, D. H., and Bi, Y. X. (2017). Research progress of wild soybean germplasm and utilization. *ACS Agric. Sci. Technol.* 18, 812–817. doi: 10.16175/j.cnki.1009-4229.2017.05.016
- Dalling, J. W., Davis, A. S., Schutte, B. J., and Arnold, A. E. (2011). Seed survival in soil: interacting effects of predation, dormancy and the soil microbial community. *J. Ecol.* 99, 89–95. doi: 10.1111/j.1365-2745.2010.01739.x
- Debeaujon, I., Léon-Kloosterziel, K. M., and Koornneef, M. (2000). Influence of the testa on seed dormancy, germination, and longevity in Arabidopsis. *Plant Physiol.* 122, 403–414. doi: 10.1104/pp.122.2.403
- Egli, D. B., TeKrony, D. M., and Heitholt, J. J. (2005). Air temperature during seed filling and soybean seed germination and vigor. *Crop Sci.* 45, 1329–1335. doi: 10.2135/cropsci2004.0029
- Ferrandis, P., Herranz, J. M., and MartinezSanchez, J. J. (1999). Effect of fire on hard-coated *Cistaceae* seed banks and its influence on techniques for quantifying seed banks. *Plant Ecol.* 144, 103–114. doi: 10.1023/A:1009816309061
- Finch-Savage, W. E., and Leubner-Metzger, G. (2006). Seed dormancy and the control of germination. *New Phytol.* 171, 501–523. doi: 10.1111/j.1469-8137.2006.01787.x
- Foley, M. E. (2001). Seed dormancy: an update on terminology, physiological genetics, and quantitative trait loci regulating germinability. *Weed Sci.* 49, 305–317. doi: 10.1614/0043-1745(2001)049[0305:SDAUOT]2.0.CO;2
- Guo, B., Lei, J. Z., and Ma, J. L. (2002). Formation of soybean hard seeds and investigation of hard seed rates of different cultivars. *Modern Agric.* 2002, 6. (in Chinese).
- Guo, J., Wang, Y., Song, C., Zhou, J., Qiu, L., Huang, H., et al. (2010). A single origin and moderate bottleneck during domestication of soybean (*Glycine max*): implications from microsatellites and nucleotide sequences. *Ann. Bot.* 106, 505–514. doi: 10.1093/aob/mcq125
- Hartwig, E. E., and Potts, H. C. (1987). Development and evaluation of impermeable seed coats for preserving soybean seed quality. *Crop Sci.* 27, 506–508. doi: 10.2135/cropsci1987.0011183X002700030015x
- Heatherly, L. G., Kenty, M. M., and Kilen, T. C. (1995). Effects of storage environment and duration on impermeable seed coat in soybean. *Field Crop Res.* 40, 57–62. doi: 10.1016/0378-4290(94)00075-N
- Heredia, A. (2003). Biophysical and biochemical characteristics of cutin a plant barrier biopolymer. *Biochim. Biophys. Acta* 1620, 1–7. doi: 10.1016/S0304-4165(02)00510-X
- Hill, H. J., West, S. H., and Hinson, K. (1986a). Soybean seed size influences expression of the impermeable seed-coat trait 1. *Crop Sci.* 26, 634–637. doi: 10.2135/cropsci1986.0011183X002600030044x
- Hill, H. J., West, S. H., and Hinson, K. (1986b). Effect of water stress during seedfill on impermeable seed expression in soybean. *Crop Sci.* 26, 807–812. doi: 10.2135/cropsci1986.0011183X002600040038x
- Huang, J., Lin, G., Xu, S., Chen, Z., and Wang, J. (1999). Water absorbing characteristics and its relation with germination in improved spring soybean varieties. *Fujian J. Agric. Sci.* 14, 8–11. (in Chinese).
- Hudson, K. (2022). Soybean protein and oil variants identified through a forward genetic screen for seed composition. *Plants* 11, 2966. doi: 10.3390/plants11212966

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Hyde, E. O. C. (1954). The function of the hilum in some Papilionaceae in relation to the ripening of the seed and the permeability of the testa. *Ann. Bot.* 11, 241–256. doi: 10.1093/oxfordjournals.aob.a083393
- Hyten, D., Song, Q., Zhu, Y., Choi, L., Nelson, R. L., Costa, J. M., et al. (2006). Impacts of genetic bottlenecks on soybean genome diversity. *Proc. Natl. Acad. Sci. U.S.A.* 103, 16666–16671. doi: 10.1073/pnas.0604379103
- Jaganathan, G. K., and Harrison, R. J. (2024). Decoding the decisive role of seed moisture content in physical dormancy break: filling the missing links. *Plant Biol. (Stuttgart Germany)* 26, 3–10. doi: 10.1111/plb.13602
- Jang, S. J., Sato, M., Sato, K., Jisuyama, Y., Fujino, K., Mori, H., et al. (2015). A single-nucleotide polymorphism in an endo-1,4- $\beta$ -Glucanase gene controls seed coat permeability in soybean. *PLoS One* 10, e0128527. doi: 10.1371/journal.pone.0128527
- Kebede, H., Smith, J. R., and Ray, J. D. (2014). Identification of a single gene for seed coat impermeability in soybean PI 594619. *Theor. Appl. Genet.* 127, 1991–2003. doi: 10.1007/s00122-014-2355-2
- Keim, P., Diers, B. W., and Shoemaker, R. C. (1990). Genetic analysis of soybean hard seediness with molecular markers. *Theor. Appl. Genet.* 79, 465–469. doi: 10.1007/BF00226154
- Keith, B. C., and Delouche, J. C. (1999). “Seed quality, production and treatment,” in *Soybean production in the midsouth*. eds. IG Heatherly and H. F. Hodges (Boca Raton, FL: CRC Press), 197–239.
- Kilen, T. C., and Hartwig, E. E. (1978). An inheritance study of impermeable seed in soybeans. *Field Crop Res.* 1, 65–70. doi: 10.1016/0378-4290(78)90007-2
- Kim, M. S., Lozano, R., Kim, J. H., Bae, D. N., Kim, S. T., Park, J. H., et al. (2021). The patterns of deleterious mutations during the domestication of soybean. *Nat. Commun.* 12, 97. doi: 10.1038/s41467-020-20337-3
- Kim, M. Y., Van, K., Kang, Y. J., Kim, K. H., and Lee, S. H. (2012). Tracing soybean domestication history: From nucleotide to genome. *Breed. Sci.* 61, 445–452. doi: 10.1270/jsbbs.61.445
- Koornneef, M., Bentsink, L., and Hilhorst, H. (2002). Seed dormancy and germination. *Curr. Opin. Plant Biol.* 5, 33–36. doi: 10.1016/S1369-5266(01)00219-9
- Kulik, M. M., and Yeklich, R. W. (1991). Soybean seed coat structures. Relationship to weathering and infection by the fungus *Phomopsis phaseoli*. *Crop Sci.* 31, 108–113. doi: 10.2135/cropsci1991.0011183X003100010027x
- Kuo, W. H. J. (1989). Delayed permeability of soybean seeds: Characteristics and screening methodology. *Seed Sci. Technol.* 27, 131–142.
- Larson, G., Piperno, D. R., Allaby, R. G., Purugganan, M. D., Andersson, L., Arroyo-Kalin, M., et al. (2014). Current perspectives and the future of domestication studies. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6139–6146. doi: 10.1073/pnas.1323964111
- Li, Y., Guan, R., Liu, Z., Ma, Y., Wang, L., Li, L., et al. (2008). Genetic structure and diversity of cultivated soybean (*Glycine max* (L.) Merr.) landraces in China. *Theor. Appl. Genet.* 117, 857–871. doi: 10.1007/s00122-008-0825-0
- Li, G. F., and Huang, W. (1994). Water absorption traits of *Glycine soja*. *Soybean Sci.* 13, 376–379. (in Chinese).
- Li, D., Zhang, Z., Gao, X., Zhang, H., Bai, D., Wang, Q., et al. (2023). The elite variations in germplasm for soybean breeding. *Mol. Breed.* 43, 37. doi: 10.1007/s11032-023-01378-0
- Liang, M., Davis, E., Gardner, D., Cai, X., and Wu, Y. (2006). Involvement of *AtLAC15* in lignin synthesis in seeds and in root elongation of *Arabidopsis*. *Planta* 224, 1185–1196. doi: 10.1007/s00425-006-0300-6
- Liang, R., Liu, J. L., Ji, X. Q., Olsen, K., Qiang, S., and Song, X. L. (2023). Fitness and hard seediness of F<sub>2</sub> and F<sub>3</sub> descendants of hybridization between herbicide-resistant *Glycine max* and *G. soja*. *Plants* 12, 3671. doi: 10.3390/plants12213671
- Liu, C., Chen, X., Wang, W., Hu, X., Han, W., He, Q., et al. (2021). Identifying wild versus cultivated gene-alleles conferring seed coat color and days to flowering in soybean. *Int. J. Mol. Sci.* 22, 1559. doi: 10.3390/ijms22041559
- Liu, B., Fujita, T., Yan, Z. H., Sakamoto, S., Xu, D., and Abe, J. (2007). QTL mapping of domestication-related traits in soybean (*Glycine max*). *Ann. Bot.* 100, 1027–1038. doi: 10.1093/aob/mcm149
- Ma, F., Cholewa, E., Mohamed, T., Peterson, C. A., and Gijzen, M. (2004). Cracks in the palisade cuticle of soybean seed coats correlate with their permeability to water. *Ann. Bot.* 94, 213–228. doi: 10.1093/aob/mch133
- McDonald, M. R., Vertucci, C. W., and Roos, E. E. (1988). Seed coat regulation of soybean seed imbibition. *Crop Sci.* 28, 987–992. doi: 10.2135/cropsci1988.0011183X002800060025x
- Medic, J., Atkinson, C., and Hurburgh, C. R. (2014). Current knowledge in soybean composition. *J. Am. Oil Chem. Soc.* 91, 363–384. doi: 10.1007/s11746-013-2407-9
- Meyer, C. J., Stuedle, E., and Peterson, C. A. (2007). Patterns and kinetics of water uptake by soybean seeds. *J. Exp. Bot.* 58, 717–732. doi: 10.1093/jxb/erl244
- Mohamed-Yasseen, Y., Barringer, S. A., and Costanza, S. S. (1994). The role of seed coats in seed viability. *TBR* 60, 426–439. doi: 10.1007/BF02857926
- Moore, S. H. (1991). Implications of the interaction between seed size and scarification on developing and conditioning hardseeded soybean. *Seed Sci. Technol.* 15, 22–28.
- Moore, S., and Yeklich, R. (2000). Using hard seed coats to preserve seed quality in soybean. *CSSA Special Publications* 6, 59–65. doi: 10.2135/cssaspecpub31.c6
- Mullin, W. J., and Xu, W. (2001). Study of soybean seed coat components and their relationship to water absorption. *J. Agric. Food Chem.* 49, 5331–5335. doi: 10.1021/jf101030s
- Nik, W. Z., and Parbery, D. G. (1977). Studies of seed-borne fungi of tropical pasture legume species. *Aust. J. Agric. Res.* 28, 821–841. doi: 10.1071/AR9770821
- Nooden, L. D., Blakley, K. A., and Grzybowski, J. M. (1985). Control of seed coat thickness and permeability in soybean: a possible adaptation to stress. *Plant Physiol.* 79, 543–545. doi: 10.1104/pp.79.2.543
- Paulsen, T. R., Colville, L., Kranner, I., Daws, M. I., Högestedt, G., Vandvik, V., et al. (2013). Physical dormancy in seeds: a game of hide and seek? *New Phytol.* 198, 496–503. doi: 10.1111/nph.12191
- Paulsen, T. R., Högestedt, G., Thompson, K., Vandvik, V., Eliassen, S., and Leishman, M. (2014). Conditions favouring hard seediness as a dispersal and predator escape strategy. *J. Ecol.* 102, 1475–1484. doi: 10.1111/1365-2745.12323
- Pietrzak, L. N., Fregeau-Reid, J., Chatson, B., and Blackwell, B. (2002). Observation on water distribution in soybean seed during hydration processes using nuclear magnetic resonance imaging. *Can. J. Plant Sci.* 82, 513–519. doi: 10.4141/P01-150
- Potts, H. C., Duangpatra, J., Hairston, W. G., and Delouche, J. C. (1978). Some influences of hardseededness on soybean seed quality. *Crop Sci.* 18, 221–224. doi: 10.2135/cropsci1978.0011183X001800020006x
- Qutob, Q. D., Ma, M. F., Peterson, P. A. C., Bernards, B. A. M., and Gijzen, G. M. (2008). Structural and permeability properties of the soybean seed coat. *Botany* 86, 219–227. doi: 10.1139/B08-002
- Ragus, L. N. (1987). Role of water absorbing capacity in soybean germination and seedling vigour. *Seed Sci. Technol.* 15, 285–296.
- Ramakrishna, G., Kaur, P., Nigam, D., Chaduvula, P. K., Yadav, S., Talukdar, A., et al. (2018). Genome-wide identification and characterization of InDels and SNPs in *Glycine max* and *Glycine soja* for contrasting seed permeability traits. *BMC Plant Biol.* 18, 141. doi: 10.1186/s12870-018-1341-2
- Ranathunge, K., Shao, S., Qutob, D., Gijzen, M., Peterson, C. A., and Bernards, M. A. (2010). Properties of the soybean seed coat cuticle change during development. *Planta* 231, 1171–1188. doi: 10.1007/s00425-010-1118-9
- Rolston, M. P. (1978). Water impermeable seed dormancy. *Bot. Rev.* 44, 365–396. doi: 10.1007/BF02957854
- Roy, K. W., Keith, B. C., and Andrews, C. H. (1995). Resistance of hardseeded soybean lines to seed infection by *Phomopsis*, other fungi and soybean mosaic virus. *Can. J. Plant Pathol.* 16, 122–128. doi: 10.1080/07060669409500769
- Saio, K. (1976). Soybeans resistant to water absorption. *Cereal Food World* 21, 168–173.
- Saio, K., Arai, K., and Watanabe, T. (1973). Fine structure of soybean seed coat and its changes on cooking. *Cereal Sci. Today* 18, 197–201.
- Sakamoto, S. I., Abe, J., Kanazawa, A., and Shimamoto, Y. (2004). Marker-assisted analysis for soybean hard seediness with isozyme and simple sequence repeat loci. *Breed. Sci.* 54, 133–139. doi: 10.1270/jsbbs.54.133
- Shao, S., Meyer, C. J., Ma, F., Peterson, C. A., and Bernards, M. A. (2007). The outermost cuticle of soybean seeds: chemical composition and function during imbibition. *J. Exp. Bot.* 58, 1071–1082. doi: 10.1093/jxb/erl268
- Sharma, S., Paul, P. J., Sameer, K. C. V., and Nimje, C. (2020). Utilizing wild *Cajanus platycarpus*, a tertiary gene pool species for enriching variability in the primary gene pool for pigeonpea improvement. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.01055
- Singh, R. K., Raipuria, P. K., Bhatia, V. S., Rani Pushpendra, A., Husain, S. M., Satyavathi, C. T., et al. (2008). Identification of SSR markers associated with seed coat permeability and electrolyte leaching in soybean. *Physiol. Mol. Biol. Pla.* 14, 173–177. doi: 10.1007/s12298-008-0016-0
- Smykal, P., Vernoud, V., Blair, M. W., Soukup, A., and Thompson, R. D. (2014). The role of the testa during development and in establishment of dormancy of the legume seed. *Front. Plant Sci.* 5. doi: 10.3389/fpls.2014.00351
- Soltani, A., Walter, K. A., Wiersma, A. T., Santiago, J. P., Quijque, M., Chitwood, D., et al. (2021). The genetics and physiology of seed dormancy, a crucial trait in common bean domestication. *BMC Plant Biol.* 21, 58–58. doi: 10.1186/s12870-021-02837-6
- Souza, F. H., and Marcos-Filho, J. (2001). The seed coat as a modulator of seed-environment relationships in Fabaceae. *Braz. J. Bot.* 24, 365–375. doi: 10.1590/S0100-84042001000400002
- Sun, L., Miao, Z., Cai, C., Zhang, D., Zhao, M., Wu, Y., et al. (2015). *GmHs1-1*, encoding a calcineurin-like protein, controls hard-seededness in soybean. *Nat. Genet.* 47, 939–943. doi: 10.1038/ng.3339
- Sun, X. M., Wang, Z., Li, S. G., Meng, F. F., Wang, S. M., and Zhang, J. P. (2014). Progress on formation mechanism and breaking methods of hard seed in soybean. *Soybean Sci. Tech.* 3, 23–27.
- Taira, H. (1990). Quality of soybeans for processed foods in Japan. *Jpn. Agric. Res. Q.* 24, 224–230.
- Takahashi, Y., Sakai, H., Ariga, H., Teramoto, S., Shimada, T. L., Eun, H., et al. (2023). Domesticating *Vigna stipulacea*: chromosome-level genome assembly reveals *VsPSATI1* as a candidate gene decreasing hard-seededness. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1119625
- Takahashi, Y., Sakai, H., Yoshitsu, Y., Muto, C., Anai, T., Pandiyan, M., et al. (2019). Domesticating *Vigna Stipulacea*: A potential legume crop with broad resistance to biotic stresses. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.01607

- Tyler, J. M. (1997). Effect of impermeable seed coat on germination of seed from early maturing of soybean. *Seed Sci. Technol.* 19, 45–50.
- Vieira, R. D., Tekrony, D., and Egli, D. B. (1992). Effect of drought and defoliation stress in the field on soybean seed germination and vigor. *Crop Sci.* 32, 023110–023110. doi: 10.2135/cropsci1992.0011183X003200020037x
- Vijayan, S. S., Nagarajappa, N., and Ranjitha, H. P. (2023). Seed coat composition in black and white soybean seeds with differential water permeability. *Plant Biol.* 25, 935–943. doi: 10.1111/plb.13551
- Wang, J. L. (1999). Study on preservation of soybean germplasm using soybean hard seed. *Soybean Sci.* 18, 351–354.
- Wang, N., Feng, S., Ma, X., Chen, Q., Liu, C., and Qi, Z. (2023). Meta-analysis and multiomics of a chromosome segment substitution line reveal candidate genes associated with seed hardness in soybean. *J. Agric. Food Chem.* 71, 16840–16854. doi: 10.1021/acs.jafc.3c03950
- Wang, M., Li, W. Z., Fang, C., Xu, F., Liu, Y. C., Wang, Z., et al. (2018). Parallel selection on a dormancy gene during domestication of crops from multiple families. *Nat. Genet.* 50, 1435–1441. doi: 10.1038/s41588-018-0229-2
- Wang, F., Sun, X., Liu, B., Kong, F., Pan, X., and Zhang, H. (2022). A polygalacturonase gene *PG031* regulates seed coat permeability with a pleiotropic effect on seed weight in soybean. *Theor. Appl. Genet.* 135, 1603–1618. doi: 10.1007/s00122-022-04057-4
- Watanabe, S., Tajuddin, T., Yamanaka, N., Hayashi, M., and Harada, K. (2004). Analysis of QTLs for reproductive development and seed quality traits in soybean using recombinant inbred lines. *Breed. Sci.* 54, 399–407. doi: 10.1270/jsbbs.54.399
- Wen, Z., Ding, Y., Zhao, T., and Gai, J. (2009). Genetic diversity and peculiarity of annual wild soybean (*G. soja* Sieb. et Zucc.) from various ecoregions in China. *Theor. Appl. Genet.* 119, 371–381. doi: 10.1007/s00122-009-1045-y
- Wen, Z., Lu, X., Wen, J., Wang, Z., and Chai, M. (2024). Physical seed dormancy in Legumes: molecular advances and perspectives. *Plants* 13, 1473. doi: 10.3390/plants13111473
- Western, T. L., Burn, J., Tan, W. L., Skinner, D. J., Martin-McCaffrey, L., Moffat, B. A., et al. (2001). Isolation and characterization of mutants defective in seed coat mucilage secretory cell development in Arabidopsis. *Plant Physiol.* 127, 998–1011. doi: 10.1104/pp.010410
- Wysmierski, P. T., and Vello, N. A. (2013). The genetic base of Brazilian soybean cultivars: evolution over time and breeding implications. *Genet. Mol. Biol.* 36, 547–555. doi: 10.1590/S1415-47572013005000041
- Xu, L., Li, J., Yin, P., Wang, G., Yan, X., and Sun, B. (2009). Testa morphology structure and germination characteristic of *Glycine soja*. *Soybean Sci.* 28, 641–646.
- Yaklich, R. W., and Cregan, P. B. (1989). Registration of Md84–2751, a soybean germplasm line with im-permeable seed coat. *Crop Sci.* 29, 835–836. doi: 10.2135/cropsci1989.0011183X002900030077x
- Yaklich, R. W., Vigil, E. L., Erbe, E. F., and Wergin, W. P. (1992). The fine structure of aleurone cells in the soybean seed coat. *Protoplasma* 167, 108–119. doi: 10.1007/BF01353586
- Yaklich, R. W., Vigil, E. L., and Wergin, W. P. (1986). Pore development and seed coat permeability in soybean. *Crop Sci.* 26, 616–624. doi: 10.2135/cropsci1986.0011183X002600030041x
- Yuan, B., Yuan, C., Wang, Y., Liu, X., Qi, G., and Wang, Y. (2022). Identification of genetic loci conferring seed coat color based on a high-density map in soybean. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.968618
- Zeng, Q. Y., Yang, C. Y., Ma, Q. B., Li, X. P., Dong, W. W., and Nian, H. (2012). Identification of wild soybean miRNAs and their target genes responsive to aluminum stress. *BMC Plant Biol.* 12, 182. doi: 10.1186/1471-2229-12-182
- Zhang, B., Chen, P., Chen, C. Y., Wang, D., Shi, A., Hou, A., et al. (2008). Quantitative trait loci mapping of seed hardness in soybean. *Crop Sci.* 48, 1341–1349. doi: 10.2135/cropsci2007.10.0544
- Zhang, B., Chen, P., Shi, A., Hou, A., Ishibashi, T., and Wang, D. (2009). Putative quantitative trait loci associated with calcium content in soybean seed. *J. Hered.* 100, 263–269. doi: 10.1093/jhered/esn096
- Zhou, S., Sekizaki, H., Yang, Z., Sawa, S., and Pan, J. (2010). Phenolics in the seed coat of wild soybean (*Glycine soja*) and their significance for seed hardness and seed germination. *J. Agric. Food Chem.* 58, 10972. doi: 10.1021/jf102694k
- Zhou, Z., Yu, J., Wang, Z., Gou, Z., Lyu, J., Li, W., et al. (2015). Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat. Biotechnol.* 33, 408–414. doi: 10.1038/nbt.3096
- Zhuang, Y. B., Li, X. M., Hu, J. M., Xu, R., and Zhang, D. J. (2022). Expanding the gene pool for soybean improvement with its wild relatives. *Abiotech* 3, 115–125. doi: 10.1007/s42994-022-00072-7