

Review

Single-Cell Omics in Legumes: Research Trends and Applications

Yaohua Li [†], Md Sabbir Hossain [†]  and Marc Libault ^{*}

Division of Plant Science and Technology, Interdisciplinary Plant Group and Christopher S. Bond Life Sciences Center, University of Missouri-Columbia, Columbia, MO 65211, USA; ylv9@missouri.edu (Y.L.); mhggr@missouri.edu (M.S.H.)

^{*} Correspondence: libaultm@missouri.edu

[†] These authors contributed equally to this work.

Abstract

Legumes are important food crops and play a central role in sustainable agriculture through their ability to form symbiosis with rhizobia, soil bacteria that fix atmospheric nitrogen. Recent advances in single-cell and spatial transcriptomics, along with single-cell epigenomics, have enabled high-resolution analysis of gene expression dynamics and the prediction of cell-type-specific regulatory networks. In this review, we highlight recent progress in the use of single-cell omics in legumes, with a particular focus on how genes functioning in distinct cell types contribute to plant development, responses to pathogens, stress-induced plasticity, and the establishment of root nodule symbioses. Case studies in *Medicago truncatula*, *Lotus japonicus*, *Glycine max*, and *Arachis hypogaea* illustrate the shift from bulk to single-cell multi-omics. We conclude by outlining current limitations and future directions for building integrated legume cell atlases that will support translational research and crop improvement.

Keywords: single-cell transcriptomics; cell identity and regulatory networks; comparative cell atlas; nitrogen fixation; synthetic biology; translational breeding; legume crops



Academic Editor: Koji Murai

Received: 31 August 2025

Revised: 23 November 2025

Accepted: 25 November 2025

Published: 27 November 2025

Citation: Li, Y.; Hossain, M.S.; Libault, M. Single-Cell Omics in Legumes: Research Trends and Applications. *Plants* **2025**, *14*, 3615. <https://doi.org/10.3390/plants14233615>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Understanding the diversity and specialized functions of the cells composing legume plants and their specific transcriptome are essential for linking gene activity to plant development, stress adaptation, and symbiotic interactions. For instance, legumes uniquely engage in root nodule symbiosis with rhizobia, forming specialized root organs that harbor nitrogen-fixing bacteroids and, at the same time, offer a tractable model to study plant cell differentiation and organogenesis [1]. Beyond their ecological significance, legumes also provide powerful models to investigate organ development, plant-microbe signaling, and adaptive plasticity under stress. Advances in single-cell and spatial transcriptomics now allow researchers to dissect cellular heterogeneity across diverse tissues, including roots, leaves, seeds, pods, and nodules at single-cell resolution [2–5]. These approaches enable cell-type-specific functional annotation and comparative genomics across species, thereby broadening the scope of legume research beyond what bulk transcriptomics can achieve.

Recent advances in single-cell and spatial transcriptomics have enabled legume research by capturing rare cell types, transient states, and mapping gene regulatory programs at high precision [2–5]. These approaches reveal cell-type diversity, developmental trajectories, and spatially organized transcriptional programs that bulk RNA-seq could not

resolve [6,7]. Applications include single-cell omics on *Lotus japonicus* (hereafter referred to as Lotus) root tips [8], various soybean organs [4,5,9], and studies of stress and pathogen responses [10] as well as root nodule symbiosis [2,11,12], including single-nucleus spatial maps [13], and functional insights [14]. Together, these studies provide a basis for redefining legume cell-type classification and enabling cross-species comparisons, while laying the groundwork for predictive modeling and translational breeding strategies.

This review synthesizes recent advances in legume single-cell biology, drawing on literature collected from PubMed and Web of Science up to August 2025 using the keywords “single-cell RNA-seq,” “spatial transcriptomics,” and the names of legume species. Peer-reviewed studies were prioritized, and recent preprints were considered when they provided new insights. It integrates multi-omics data from model and crop legumes, including *Medicago truncatula* (hereafter referred to as Medicago), Lotus, *Glycine max* (soybean), and *Arachis hypogaea* (peanut), as well as spatial transcriptomics studies. These representative legumes were selected as models to study biological nitrogen fixation, and because of their substantial agronomical value as major suppliers of plant-derived oil and protein. Specifically, soybean meal’s dry mass contains 40 to 49% of proteins, significantly contributing to the share of proteins used in global livestock feed [15]. Besides, soybean is the second most productive vegetable oil crop after oil palm, generating approximately 68 million tons of oil annually and accounting for nearly 28% of the global vegetable oil market, despite its seeds containing only about 20% oil [16]. This review highlights how these approaches have redefined cell-type classification, developmental trajectories, and gene regulatory programs. It also outlines methodological considerations and future priorities for cross-legume species integration, epigenomic context, and functional validation aimed at legume improvement.

2. From Bulk to Single-Cell Resolution: Building Cell-Level Insights

2.1. Findings Revealed by Bulk RNA-seq

Transcriptomic technologies have transformed plant biology over the past two decades. Early microarray studies [17] provided the first large-scale gene expression profiles including for the model legume Medicago [18] followed by RNA sequencing (RNA-seq) [19], which rapidly became the standard due to its greater sensitivity and resolution [20–22]. This technology provides a genome-wide view of gene expression from a biological sample, often from entire organs. By identifying genes that are up- or down-regulated in response to developmental cues or environmental stimuli, RNA-seq analyses uncover the role of plant genes that drive biological processes. As a result, RNA-seq has been widely used in legumes to study genetic programs controlling development, nutrient composition, biotic and abiotic stress responses, and symbiotic nitrogen fixation [23]. The first RNA-seq study in *Glycine max*, published in 2010, profiled root hair-specific transcriptomes following *Bradyrhizobium japonicum* infection and revealed rapid, genome-wide transcriptional reprogramming in root hair cells upon bacterial contact [24]. Since then, advances in sequencing technologies and decreases in cost have enabled the generation of numerous transcriptomic studies across diverse legume organs, which collectively uncovered the transcriptional reprogramming of a wide range of biological processes, including nodulation, organ development, and response to biotic and abiotic stress (e.g., drought, flooding, salinity, heat) [20,21,25–33]. For instance, Xu et al. [33] identified strong induction of ABA-responsive transcription factors (TFs), ROS-scavenging systems, and osmoprotectant biosynthesis pathways during drought, whereas Liu et al. [29] revealed substantial transcriptional reprogramming under salt stress affecting photosynthesis, ion transport, and phytohormone signaling pathways. Bulk RNA-seq has also yielded important insights into seed biology. For instance, transcriptomes generated across soybean seed developmen-

tal stages revealed the coordinated induction of sucrose transporters, lipid biosynthesis enzymes, and storage protein genes, helping define the molecular transitions underlying seed filling and composition [21,30]. Similarly, numerous bulk RNA-seq studies have been conducted in *Medicago*, *Lotus*, peanut, *Pisum sativum* L., and *Cicer arietinum* L. to identify key genes, transcriptional regulators, and adaptive mechanisms associated with tissue development, symbiotic interactions, and responses to biotic and abiotic stresses [34–46]. These datasets have greatly expanded the gene expression resources available for legumes, leading to the development of specialized online platforms such as SoyBase [47], Legume-Info [48], Lotus Base [49], MtExpress [50], and PeanutBase [51] that provide centralized access to legume transcriptomic data.

2.2. Limitations of Bulk RNA-seq in Resolving Cellular Heterogeneity

Despite its value, bulk RNA-seq only captures average gene expression across heterogeneous cell populations, lacking the resolution to dissect transcriptional activity at the single-cell level. Given that plant tissues comprise diverse cell types, each with distinct transcriptional programs, there has been a growing need for technologies capable of resolving gene expression in individual cells and cell to cell communication. To bridge this gap, early approaches isolated specific plant cell populations using techniques such as fluorescence-activated cell sorting (FACS) and laser capture microdissection (LCM), followed by microarray or bulk RNA-seq analysis. LCM-based transcriptomics has also been successfully applied in legumes to study gene expression in specific tissues [52]. For example, in *Medicago*, Limpens et al. [53] and Roux et al. [54] utilized LCM to resolve gene expression in spatial zones of nodules, shedding light on nodule development and function. Bulk RNA-seq fails to capture the expression dynamics of individual cells, particularly rare cell types, spatial and temporal variation, and low-abundance transcripts such as TFs. To support the benefit of snRNA-seq, we recently reported that diverse nodule cell types (cortical, infected, and vascular) exhibit distinct transcriptional programs that bulk RNA-seq cannot resolve [2]. Similarly, bulk approaches fail to capture dynamic transitions among sub-cell types, such as the functional differentiation of three endosperm subpopulations described by Zhang et al. [4]. Another major limitation of bulk RNA-seq is its inability to reveal the transcriptome of rare cell types. For example, in *Arabidopsis thaliana* (hereafter referred to as *Arabidopsis*) the transcriptome of the quiescent center (QC) cells, a cell type that constitutes only a very small fraction of root cells, was established using a single-cell RNA-seq approach [55]. Such a result could not be achieved using a bulk RNA-seq strategy, where the expression profiles of QC cells are largely masked by the transcriptome of other cell types. In summary, although bulk RNA-seq has generated foundational knowledge in legume biology and continues to be a valuable tool, it falls short in revealing cell-type-specific transcriptional programs and those associated with rare or transient cell states, spatial heterogeneity, and dynamic developmental transitions. These limitations underscore the need for single-cell RNA sequencing technologies, which enable transcriptomic profiling at the single-cell level and allow researchers to uncover the cellular diversity and regulatory complexity hidden within bulk measurements.

2.3. Timeline of the Use of Single-Cell/Single-Nucleus RNA-seq in Legume Research

Today, the advent of droplet-based microfluidic single-cell/single-nucleus RNA sequencing (sc/snRNA-seq) technologies has revolutionized plant transcriptomics by enabling the profiling of thousands of individual cells in a single experiment at reduced cost. Unless otherwise specified, “scRNA-seq” hereafter broadly refers to both sc/snRNA-seq approaches. This approach utilizes cell barcoding and unique molecular identifiers to accurately assign transcripts to their respective cells of origin, while minimizing amplification

biases. First demonstrated in *Arabidopsis* roots [55,56], droplet-based workflows have since been applied to various plant species and tissues, facilitating cell clustering, trajectory inference, and gene regulatory network reconstruction. In legumes, the first single-cell transcriptome was conducted in peanut leaf, identifying eight cell clusters with distinct transcriptional regulation related to leaf development [57]. Since then, scRNA-seq has been increasingly applied to legume species, including soybean, *Medicago*, peanut, and *Lotus*, covering both the early stages of rhizobial infection as well as analyses across different organs [2,4,5,8–14,24,52,53,57–65] (Figure 1). These advances have rapidly expanded legume transcriptomics research, providing cell-resolved insights into gene expression and cell biology. Beyond profiling the transcriptome of diverse and rare cell types, sc/snRNA-seq reveals pseudo-temporal changes in gene activity during development and stress responses. For example, *Medicago* nodules comprise meristematic, infection, nitrogen-fixation, and senescence zones [66], each characterized by distinct gene expression programs, unique chromatin organization, and histone mark status essential for their specific functions [54]. The emergence of scRNA-seq technologies has facilitated access to single-cell and cell-type transcriptomic resolution, enabling the precise capture of expression patterns of individual cells and cell types, including populations of infected and uninfected cells within *Medicago* nodules [59]. To further estimate the dynamic and continuous transcriptomic changes occurring in response to environmental and developmental cues, scRNA-seq enables the computational construction of cellular “trajectories” as cells transition between developmental states (e.g., during soybean embryogenesis, parenchyma cell differentiation in the peanut shell, and epidermal cell differentiation in peanut) [4,62]. Such analysis reveals transient populations, a more precise timing of gene regulation that drives dynamic processes and supports a more accurate inference of gene regulatory networks (GRNs). The latter is an important feature of scRNA-seq analysis, considering that bulk datasets may generate false positives (e.g., genes appear co-expressed but are active in different cell types) and false negatives (e.g., regulatory relationships are masked due to low expression in a small subset of cells).

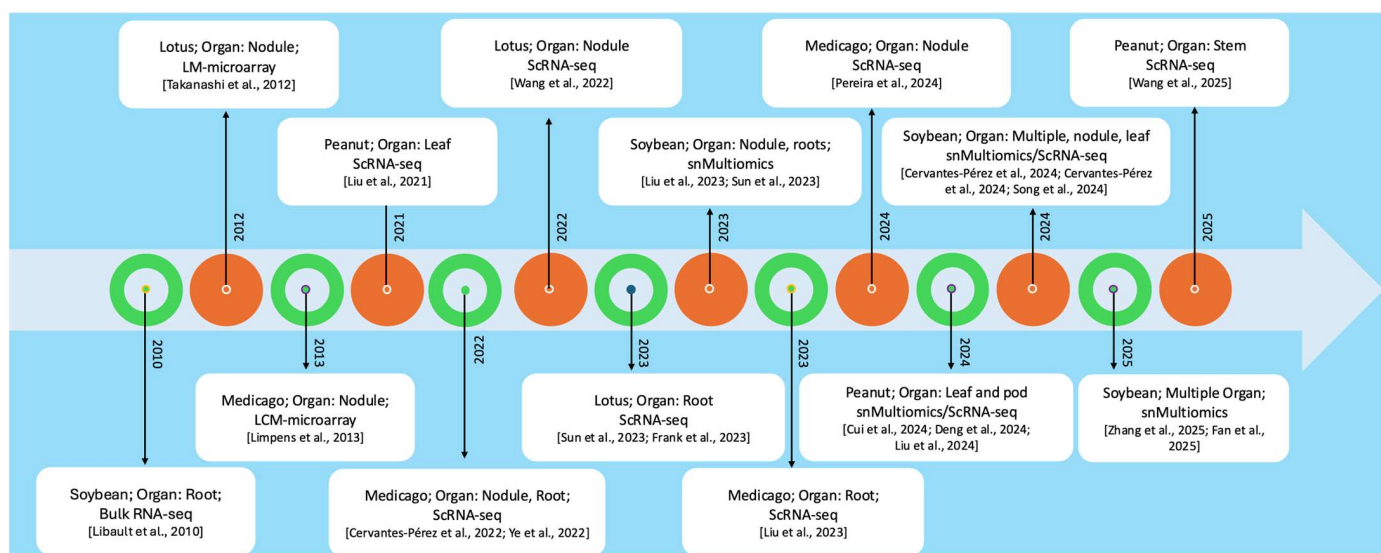


Figure 1. Timeline of key milestones towards single-cell transcriptomics research in legumes. (Note: References in this figure follow the numbering sequence of their first appearance in the main text to maintain consistency with the journal’s citation style) [2,4,5,8–14,24,52,53,57–65].

2.4. Implementation of Single-Cell Multi-omics Strategies to Predict Gene Regulatory Networks

Single-cell transcriptomic technologies have significantly advanced our understanding of cellular diversity, yet they offer only a partial view of the complex regulatory mecha-

nisms that define cellular states, which represent the dynamic variations within a single cell type driven by developmental cues or environmental conditions [67]. An integrative approach that captures multiple layers of information is needed to gain a more comprehensive understanding of plant cell biology, including their cell states. Combining scRNA-seq with single-nucleus Assay for Transposase-Accessible Chromatin using sequencing (snATAC-seq) enables simultaneous analysis of gene expression and chromatin accessibility at single-cell resolution, offering valuable insights into the epigenetic landscape that governs transcriptional regulation. Using this approach, approximately one-third of accessible chromatin regions (ACRs) have been identified as cell-type-specific, including many linked to phenotypic variations in Arabidopsis, maize, and soybean [4,68,69]. These findings underscore the potential of chromatin accessibility profiles as robust molecular markers of cell identity, comparable to or even more stable than gene expression signatures [70]. For example, in peanut, multi-omics analysis revealed the TFs interaction networks that regulate mesophyll and epidermal-guard cell differentiation. This study uncovered a core TF-fatty acid pathway module that modulates epidermal and guard cell identity, providing novel insights into epidermal cell differentiation. Notably, the expression of *WIP2* and *AP2* exhibited divergent patterns during epidermal expansion, while *KAN4* activity was found to respond to the oil synthesis pathway, mediating guard cell development [64]. In another peanut study, ATAC-seq identified 56 distinct vascular tissues-associated DNA motifs connected with pod development. Among them, aerial pegs were enriched for PIF/SGR motifs, whereas subterranean pegs showed HY5, HYH, and PIF1 motif enrichment. These light-responsive motifs highlight light-regulated transcriptional control as a key driver of pod vascular development. In the mature peanut pod, 36 distinct DNA motifs were identified, including 24 classified as ERF-related. This strong enrichment of ERF motifs suggests their key importance in driving the development and enlargement of the subterranean peanut pod [62]. In soybean, 13 sucrose transporter genes containing the DOF11 DNA-binding motif are co-upregulated during late peripheral endosperm development, a regulatory pattern that may likewise influence seed size and oil content. In soybeans, both molecular features (i.e., RNA- and ATAC-seq) have been independently captured at the single-cell level from various organs prior to their integration [4]. This approach necessarily triggers batch effects between independently processed samples. A solution to overcome this problem involves applying cell multi-omics technology, an approach that simultaneously captures both molecular layers from the same cell, linking chromatin states to transcriptional output and ultimately shaping cell behavior and fate decisions. Together, these integrative strategies support the identification of key regulatory elements and TFs that govern specific cellular states, ultimately advancing plant developmental biology and enabling precision strategies for crop improvement.

2.5. Integration of Spatial and Single-Cell Data

While single-cell transcriptomics has greatly advanced our understanding of cellular complexity, it cannot fully capture cell-to-cell interactions and tissue organization, and therefore, spatially organized biological processes involving different cell types that are fundamental to plant development, stress responses, and biotic interactions. Another limitation in non-model species, including legumes, is the limited availability of validated cell-type marker genes to support the accurate annotation of cell types based on their transcriptomic profiles. Homology-based methods for identifying markers often encounter difficulties due to gene duplication, neofunctionalization, or lineage-specific gene loss, which can lead to potential inaccuracies in defining cell types. To overcome these constraints, spatial transcriptomics has emerged as a complementary approach to scRNA-seq technologies. By preserving the spatial organization of tissues, spatial transcriptomic techniques enable the

mapping of gene expression directly onto histological structures. Spatial transcriptomics can be categorized into (1) RNA capture-based approaches (e.g., 10× Visium, Slide-seq) where tissue sections are placed on slides patterned with spatially barcoded primers that capture RNA molecules and tag them to record their spatial locations; (2) in probe- or imaging-based approaches [e.g., multiplexed error-robust fluorescence in situ hybridization (MERFISH), sequential FISH (seqFISH), Xenium in situ], where fluorescently labeled probes hybridize target transcripts on a tissue cross-section, enabling direct visualization of transcripts locations at high spatial resolution [71]. It allows more precise characterization of tissue diversity, cell identity, and spatially regulated gene expression, offering insights into how transcriptional activity is influenced by cellular neighborhoods and microenvironmental context [2,4]. Recent studies in soybean have demonstrated the power of integrating spatial transcriptomics with scRNA-seq to profile cell-type-specific transcriptional activity within a spatial framework. Spatial RNA-seq across five tissue types, including multiple developmental stages of seeds, not only enabled accurate cell-type annotation but also revealed spatially differentiated sub-cell types [4]. In another study, the combination of Stereo-seq and scRNA-seq identified novel genes associated with nodule development in soybean [5]. Furthermore, integrating scRNA-seq with high-resolution spatial transcriptomic mapping of mature nodules revealed distinct transcriptomic signatures and uncovered the coexistence of multiple *Bradyrhizobium diazoefficiens*-infected cell subpopulations within mature nodules [2]. These advances highlight the potential of spatially resolved multi-omics to deepen our understanding of plant biology and inform targeted crop improvement strategies.

3. Developing Single-Cell Legume Atlases to Capture Cellular Diversity and Plant Cell Environmental Responses

3.1. Developmental Atlases of Plant Organs: Single-Cell and Spatial Transcriptomes in Legumes

Large-scale single-cell transcriptomic profiling across multiple organs provides a comprehensive understanding of organ-level transcriptional dynamics, reveals novel genetic programs including those conserved among different cell types, and supports the functional characterization of plant genes. Advances in single-cell and spatial transcriptomics have enabled the construction of developmental atlases for non-model crops, including legumes, providing cellular-resolution insights into tissue complexity. For instance, a recent single-nucleus transcriptome atlas of peanut constructed from leaves, stems, hypocotyls, and roots identified distinct cell types, reconstructed developmental trajectories, and highlighted key TFs and interaction networks regulating hypocotyl cell differentiation. Among these TFs, *AhWRKY70* was identified as a potential regulator of stem elongation acting by modulating the auxin and ethylene signaling pathways [65]. In a complementary study by Cui et al. [62] profiled the nuclear transcriptome across different stages of peanut pod development, revealing gravity and light responsive genes involved in pod formation and delineating differentiation trajectories of parenchyma cells. To complement these studies, leaf scRNA-seq isolated from peanut seedlings grown under dark and light conditions uncovered gene networks underlying light-mediated repression of cell division and hormonal regulation of epidermal cell differentiation [63]. Another scRNA-seq study on peanut leaf blades revealed distinct transcriptional states of mesophyll and epidermal cells, suggesting that palisade mesophyll cells may transition into spongy cells and that epidermal cells emerge earlier than primordium cells [57].

In soybean, Zhang et al. [4] recently developed a spatially resolved multi-omic single-cell atlas by integrating scRNA-seq, snATAC-seq, and spatial transcriptomics across various organs, including leaf, hypocotyl, root, nodule, young pod, and seeds. This study identified three transcriptionally distinct sub-cell types in the endosperm and 13 sucrose transporter

genes, including *GmSWEET15a* and *GmSWEET10a*, which were co-expressed in late peripheral endosperm and regulated by *GmDOF11* [4]. In another study, researchers developed a whole-soybean single-cell transcriptome atlas, identifying distinct cell clusters across 10 organs by integrating spatial transcriptomics and orthologous gene relationships. TF analysis revealed that co-expressed TF modules are sufficient to define the identities of organs and cells. For example, within the *GmWRKY* family, 20 genes were preferentially expressed in root epidermal clusters (mainly clusters 1, 2, and 4 excluding root hairs) and cortical tissues, while 33 genes showed predominant expression in the root epidermal (cluster 4) and root hair cells (cluster 3) [9]. To complement these resources, Fan et al. [5] developed a scRNA-seq atlas across five soybean organs including the root tip, shoot apex, leaf, stem, and root nodule that led to the characterization of a distinct stem cell niche cluster marked by the expression of *GmPLT4*, *GmPLT1a*, and *GmPLT1b*, which belong to the *GmAP2* TF family and are closely related to *AtPLETHORA* genes, master regulators of root stem cell activity in *Arabidopsis* [72]. Together, these atlases provide reference frameworks for dissecting transcriptional and epigenetic regulation of organ development in legumes, enabling comparative genomics across species and guiding targeted crop improvement strategies (Figure 2).

3.2. Stress and Immunity-Responsive Atlases

Biotic and abiotic environmental factors, such as pathogen attack, drought, heat, and nutrient deficiency, profoundly influence plant growth and productivity. They trigger complex signaling cascades and transcriptional reprogramming. These responses vary between tissues and cell types, reflecting specialized cellular strategies for adaptation [73–76]. At the initial site of pathogen entry, epidermal and root hair cells rapidly activate pattern-triggered immunity (PTI) through receptors such as *FLS2* and *EFR*, initiating MAP kinase cascades and early defense gene expression [77]. Guard cells also contribute to modulating abscisic acid (ABA) signaling, which induces stomatal closure as a barrier against fungal entry. Meanwhile, epidermal cells upregulate *MYB122* to promote indole glucosinolate production. Inner tissues, including cortical and vascular cells, mount subsequent immune responses involving salicylic acid (SA)-mediated signaling, bursts of reactive oxygen species (ROS), and activation of pathogenesis-related (PR) genes while vascular cells are enriched in nucleotide-binding leucine-rich repeat (NLR) receptors for intracellular immune surveillance [78]. The advent of scRNA-seq technologies has enabled single-cell resolution in dissecting these heterogeneous biotic responses [79–81]. For instance, recent studies have used these approaches to uncover cell-type-specific responses to biotic stresses in *Arabidopsis* leaves, such as infections by *Pseudomonas* and *Colletotrichum* [74,76,78]. Additionally, specialized immune-active cells, such as primary immune responder (PRIMER) cells, have been identified as local hubs for pathogen detection that relay early immune signals to neighboring “bystander” cells, facilitating long-distance defense communication [82]. Other studies in *Arabidopsis* reported that only subsets of cortical or mesophyll cells strongly activate NLR receptors, defense-related TFs, and genes associated with antimicrobial metabolites, whereas the same type of neighboring cells remain transcriptionally inactive [78,82]. Similar cellular heterogeneity has been observed in *Arabidopsis* roots, where scRNA-seq analyses have revealed cell-type-specific immune responses, including the induction of triterpenoid biosynthesis pathways in epidermis and procambium cells to balance beneficial and pathogenic microbial interactions [83]. Taken together, these studies on the model species *Arabidopsis* support the relevance of the use of single cell omics approach to better decipher the programs used by different cell types to respond to environmental stresses.

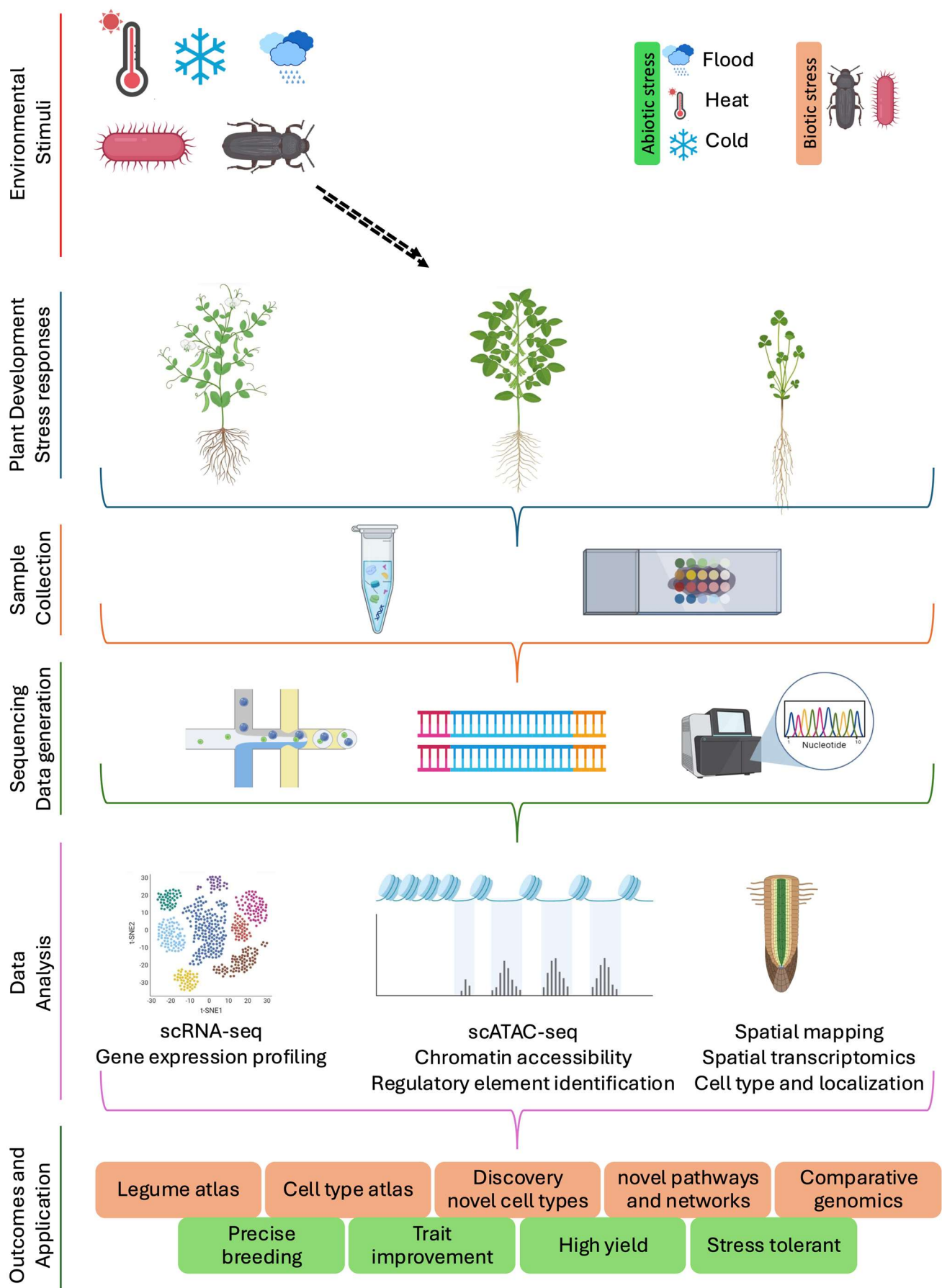


Figure 2. Integrated multi-omics approaches for a comprehensive understanding of plant growth, development, and environmental responses. Samples from different tissues, developmental stages, or

stress conditions are processed to generate scRNA-seq, snATAC-seq and spatial transcriptomics data. These datasets are integrated to construct a legume cell atlas that captures gene expression profiles, cell type localization, and regulatory landscapes. Cross-species comparisons of the integrated atlas provide functional and evolutionary insights into legume biology. Collectively, these approaches enable the identification of key regulatory genes and pathways, contributing to the development of resilient crop varieties. The arrow indicates the environmental signal responded by plants. The orange box highlights applications while green box highlights the future perspectives of single cell transcriptomics. Created with BioRender.com.

When considering legumes, in soybean, scRNA-seq has been employed to examine the transcriptional responses of soybean leaf cells to Soybean Mosaic Virus (SMV) infection. The study analyzed the transcriptome of 50,294 leaf protoplasts from infected and control plants and identified 19 distinct cell clusters, providing a high-resolution map of cell-type-specific transcriptional dynamics [10]. The *tau-class glutathione S-transferase* (GSTUs) are key mediators of plant responses to diverse abiotic and biotic stresses, including viral infections. Among the 57 *GSTU* genes, *GmGSTU23* and *GmGSTU24* were strongly upregulated in epidermal, mesophyll, and vascular cells upon SMV infection in soybean. Functional assays in *Nicotiana benthamiana* confirmed their antiviral activity, highlighting these GSTUs as key contributors to soybean defense mechanisms [10]. Of course, mutualistic symbiotic interactions between legumes and soil microbes (i.e., mycorrhization and nodulation) have also been subject to a significant number of scRNA-seq studies (see below).

Although single-cell atlases have revealed extensive cellular diversity in legumes, future research should integrate developmental and stress-responsive datasets to uncover cell states transition. Key questions include: how do different cell types of coordinates signaling and stress responses? What is the level of conservation of these regulatory mechanisms and the role of associated TFs across legumes and other plant species? What is the impact of chromatin accessibility modulating these processes? Moving beyond descriptive cataloguing, legume single-cell research is now positioned to develop predictive models that connect cell-type-specific gene regulation with plant adaptation, ultimately guiding the engineering of stress-resilient and high-yielding crops.

4. Symbiosis as a Model: Insights from Nodulation and Mycorrhizal Interactions

4.1. Root Nodule Symbiosis: As a Model for Organogenesis and Signaling

Legume nodulation results from the mutualistic symbiotic interaction between plants (mostly from the legume clade) and soil microbes collectively referred to as rhizobia (e.g., *Rhizobium*, *Bradyrhizobium*, *Sinorhizobium*). This process resulted in the formation of nitrogen-fixing nodules. Under low-nitrogen conditions, legumes secrete flavonoids that attract rhizobia. In response, rhizobia release Nod factors (NFs), which are perceived by the root epidermal cells, triggering the curling of root hairs to entrap the bacteria and initiating the formation of an infection thread that enables the infection of the deeper cellular layers of the root. Concurrently, outer cortical cells located near the protoxylem, particularly those corresponding to cortex layer 4, undergo dedifferentiation and cell cycle reactivation to form a nodule primordium, ultimately developing into a mature nodule [84,85]. Nodule organogenesis is an excellent model for studying plant organ development, as it demonstrates how differentiated root cells can re-enter the cell cycle, change identity, and acquire specialized functions. This developmental reprogramming is tightly coordinated by bacterial signals, plant hormones such as cytokinin and auxin gradients, and a complex network of TFs that collectively regulate cortical cell dedifferentiation, primordium initiation, and the spatial organization of mature nodules [12]. Because nodulation involves

diverse cell lineages including root hair cells, dividing cortical and pericycle cells, and infected and uninfected nodule cells, each with distinct transcriptional programs. Single-cell transcriptomics has emerged as a powerful tool for dissecting the dedifferentiation and redifferentiation trajectories, and the discovery of novel regulatory genes controlling nodule development.

Focusing on early symbiotic events, Pereira et al. [12] demonstrated that *Medicago* epidermal root hair and stele sub-cell types participate in symbiotic responses during infection and in regulating nodule proliferation. Consistently, *Medicago* root hair, cortex, endodermis, and pericycle cells exhibited the strongest differential gene regulation after 48 h of rhizobium inoculation [58], which aligns with global gene expression reprogramming observed in the epidermis and cortex within 30 min of NFs treatment [61], indicating that early infection triggers a multilayered cascade that propagates from the epidermis toward inner cortical tissues. In *Lotus*, Frank et al. [14] identified and profiled several populations of cells at different stages of the infection process, including infected root hair and cortical cells, nodule cells, and bacteroid-containing cells [14]. Another scRNA-seq study of *Lotus* roots identified seven major cell types and revealed associated regulatory programs, including those for phytohormones and nodulation-related gene, showing that species-specific nodulation programs still converge on conserved hormone-responsive and infection-responsive pathways across legumes [8].

Later, in soybean nodule, Zhang et al. [4] highlighted the conserved gene regulatory networks underlying soybean-rhizobium symbiosis [4]. Focusing on metabolic and developmental mechanisms, Sun et al. [60] reported that soybean purine biosynthesis genes were differentially expressed across specific nodule cell types, including infected cells, inner cortex, and phloem, highlighting coordinated intercellular metabolic partitioning of the purine biosynthesis pathway. The study also identified crucial roles for cytokinin and its receptor *GmCRE1*, which are essential for proper nodule formation and efficient nitrogen fixation [60]. These results collectively illustrate how metabolic specialization emerges across infected and uninfected domains, reinforcing the view that nodules function as spatially integrated metabolic organs rather than homogeneous structures. To complement these studies, single-cell transcriptomic maps of soybean roots and mature nodules revealed multiple subpopulations of rhizobia-infected cells, including actively nitrogen-fixing and senescing cells. Gene co-expression analyses also revealed the role of specific TFs and regulatory modules to control the infection, nitrogen fixation, and senescence programs in these subpopulations [2]. Similarly, uninfected cells in the soybean nodule specialize into several distinct subgroups which are essential to support the symbiosis [13]. Likewise, single-cell transcriptomic analysis in *Medicago* revealed that uninfected cells in the nitrogen fixation zone contribute to nitrogen assimilation through asparagine biosynthesis [59]. Similarly, a study in *Medicago* identified rare cell subtypes with distinct roles in nodule development, showing that uninfected cells differentiate into specialized subtypes mediating nutrient and energy exchange for efficient symbiotic nitrogen fixation, while infected cells comprise subtypes enriched in infection thread and symbiotic nitrogen fixation genes, highlighting their critical roles in rhizobial infection and nodule development [61]. Another single-cell analysis in *Medicago* also revealed unique transcriptional signatures for each cell type, with extensive intercellular coordination underlying successful nodule initiation [12]. Together, these studies demonstrate that single-cell omics has shifted the nodulation field from describing organ-level events to resolving the fine-scale division of labor among cell types and subtypes, thereby establishing RNS as a premier model for understanding how cellular diversification and intercellular signaling drive organogenesis and plant–microbe interactions.

4.2. Root Nodule Development: Lineage Fate and Spatial Patterning

Spatial patterning of nodules is controlled by cell-type-specific gene regulatory networks. Recent single-cell transcriptomics studies have provided single-cell resolution of lineage fate specification during nodulation. In *Medicago*, trajectory analysis revealed that nodule organogenesis originates from third-layer cortical cells, which transition through cytokinin-responsive states before adopting auxin-enriched transcriptional identities that drive primordium initiation [12]. This transition is followed by activation of key nodulation genes (e.g., *NIN*, *NF-YA1*), coupled with the expression of cytokinin degradation genes, marking a shift from cytokinin-dependent initiation to auxin-mediated primordium growth and meristem specification [12]. Another study revealed that specific subtypes of epidermal and cortical cells in *Medicago* undergo transcriptional reprogramming in response to NF signaling [61]. Similarly, distinct epidermal and cortical cell subtypes in *Lotus* roots respond to rhizobia and initiate infection thread formation [14]. In soybean nodules, infected and uninfected cells further diversify into functionally specialized subpopulations that play crucial roles in nutrient exchange and the establishment of effective symbiosis [13]. Future applications of single-cell technologies could be employed to investigate the programs governing the senescence of nodule cells, enabling reconstruction of gene regulatory networks that govern fate transitions during nodule aging and termination of symbiotic nitrogen fixation, and comparing the level of conservation of these programs between legume species; especially when considering the differential developmental patterns of determinate vs. indeterminate nodules.

4.3. Early Signaling Genes Shared in Mycorrhizal and Rhizobia Symbiosis

The establishment of both rhizobial nodulation and arbuscular mycorrhizal (AM) symbiosis relies on a conserved signaling cascade, the Common Symbiotic Signaling Pathway (CSSP) [86,87]. This pathway, originally associated with AM symbiosis, was later co-opted by legumes to support rhizobial endosymbiosis [88]. Key components of this pathway include receptor kinases, nuclear calcium channels, and calcium/calmodulin-dependent protein kinases, which ultimately activate transcriptional regulators guiding either nodule organogenesis or arbuscule formation. Comprehensive overviews of CSSP architecture and function are available in several reviews, but these frameworks were largely derived from bulk analyses and therefore could not resolve which specific epidermal or cortical populations deploy conserved versus symbiosis-specific modules [87,89,90]. Recent scRNA-seq studies now enable the dissection of symbiotic gene activity at cellular level. For instance, *MtVPY* required for infection thread progression and arbuscule development in *Medicago*, shows strong co-expression with other endosymbiotic genes in colonized cells. This dual requirement for *VPY* in both infection-thread formation and arbuscule maturation directly supports the idea that early downstream components of CSSP are functionally reused across AMS and RNS rather than operating as two independent modules. Similarly, *MtEXO70i* is specifically expressed in *Medicago* infected cortical cells, where it regulates the development of infection threads in nodulation and arbuscule branching in arbuscular mycorrhizal symbiosis [88]. Because *EXO70i* operates only in colonized cortical lineages, its expression pattern provides cellular-level evidence that CSSP-mediated signaling bifurcates downstream into symbiosis-specific trafficking pathways despite sharing upstream calcium decoding. According to Liu et al., the strong overlap between genes induced by Myc-LCOs and NFs indicates that NFs can engage AMS signaling components through the common symbiotic pathway [61]. This transcriptional overlap is mechanistically consistent with CSSP-controlled nuclear Ca^{2+} oscillations being a shared activation node, explaining why single-cell profiles from LCO- and NF-treated roots converge on similar cortical subtypes. Therefore, comparative single-cell analyses of roots colonized by rhizobia and

arbuscular mycorrhizal fungi within the same host plant have strong potential to uncover cell-type-specific functions and to delineate both conserved and unique regulatory modules underlying these symbioses.

Recent single-cell studies have provided unprecedented resolution of symbiotic processes in legumes; however, many aspects of nodule biology remain unexplored. Key questions include: What is the level of coordination of cell-type-specific programs to support nodule development, bacterial infection, and plant-microbe symbiosis? Can we engineer the genome of rhizobia-infected cells to enhance their nitrogen fixation rate or lifetime? How do TFs and the accessibility of cis-regulatory elements contribute to the regulation of the genetic programs of the nodule cells? It also remains unclear how CSSP is spatially deployed during rhizobial and arbuscular mycorrhizal colonization, as well as how these pathways intersect at the single-cell level. Future research should integrate comparative single-cell omics with chromatin accessibility and spatial datasets, complemented by targeted perturbations such as cell-type-specific CRISPR, to reconstruct developmental trajectories and identify conserved versus species-specific regulatory modules.

5. Implementing Functional Genomic Solutions at Cellular Resolution

With single-cell and spatial transcriptomic atlases now available for several legumes, the next challenge is to implement functional genomic strategies at cellular resolution to enhance specific traits. These approaches aim to connect gene expression profiles with chromatin accessibility, regulatory element activity, and TF binding. By integrating high-throughput single-cell omics with perturbation assays and computational modeling, researchers can link descriptive cell atlases to causal inference and predictive modeling of cell identity programs [2,4,14,59]. When integrated with perturbation assays and computational network reconstruction, single-cell data increasingly link molecular signatures to experimentally validated regulatory circuits [12,60].

5.1. Deciphering Cell-Type-Specific Gene Expression Landscapes

Single-cell and spatial transcriptomic technologies have enabled the resolution of cell-type-specific expression patterns [55,70,91], frameworks first established in Arabidopsis and maize and now extended to legumes. In soybean, single-cell and spatial transcriptomic atlases resolved distinct root and nodule populations and identified sucrose transporters such as *GmSWEET15a* and *GmSWEET10a* as regulators of seed endosperm development [2,4]. The ability to localize these transporters to specific cell types clarified their functional role in a way that an organ-level profiling could not, given the complex mixture of cell types composing the seed and its endosperm. Work in Lotus also illustrates how transcriptional differences between neighboring cortical cells can foreshadow developmental outcomes. Cellular heterogeneity detected in infection-associated cell clusters [14] later aligned with genetic evidence showing that *LjSYMRKL1* is required for infection thread progression. The correspondence between atlas-resolved cellular states and a precise developmental requirement provides one of the clearest examples of how single-cell data can identify points of vulnerability or regulation within the symbiotic program [14]. Studies in peanut extended this idea in a different direction. Cell-resolved datasets from pods and leaves uncovered gene expression modules shaped by positional and mechanical cues, including responses to light exposure and gravity [31,32]. The spatial separation of these modules across tissues indicates that organ-specific physical environments play a major role in defining cell identity in legumes [62,63]. Importantly, several of these candidate regulators have now been validated experimentally, including soybean sucrose transporters through physiological assays and *LjSYMRKL1* through mutational analysis [4,14]. These studies reveal a common trend: legume single-cell atlases are moving beyond cell-type identification and toward

functional dissection of regulatory genes that coordinate nutrient allocation, infection, and organ development. Collectively, these studies illustrate how legume single-cell resources advance from descriptive expression maps to functional genomic frameworks, laying the groundwork for causal inference, predictive modeling, and translational applications in crop improvement.

5.2. Chromatin Accessibility and Protein–DNA Interaction Mapping

Cell identity is determined by transcriptional activity and the accessibility of cis-regulatory elements that control TF binding. Studies in *Arabidopsis* root single cells have demonstrated that cell-type-specific ACRs strongly correlate with expression profiles [70]. This observation was confirmed in a recent study on soybean through the integration of scRNA-seq and snATAC-seq to construct a spatially resolved multi-omics atlas encompassing 316,358 cells across ten organs [4]. Within nodule-infected cells, the study identified 303,199 ACRs, with approximately 40% displaying cell-type specificity. These regions were enriched for DNA binding motifs of key symbiotic regulators, including *NIN* and *NF-YA1* [92], which have functionally conserved roles in nodule organogenesis and the establishment of a successful symbiosis [93,94]. In soybean, the *GmNN1/FT2a–GmNFYA-C* module also regulates nodulation via shoot-to-root signaling [95], and recent multi-omic mapping in this species further supports TF motif enrichment/localization [4]. The convergence between spatial ATAC-defined ACRs and DAP-seq-defined TF binding sites suggests that long-distance signaling modules, such as *FT-NF-YA*, act by modulating chromatin states at infection and cortical reprogramming loci, providing a mechanistic link between systemic cues and local cell-type specification. By linking regulatory element accessibility to key TFs in infection and cortical cells, these studies create a foundation to engineer cell-type-specific regulatory circuits in legume roots and nodules. These efforts include integrating scRNA-seq and snATAC-seq profiles with TF binding datasets from DAP-seq (e.g., 148 soybean TFs [96]) or ChIP-seq, which will allow more precise mapping of TF–DNA interactions. Moving forward, cell-type-resolved chromatin accessibility assays combined with spatial ATAC-seq and DAP-seq datasets will be instrumental in validating predicted TF–DNA interactions in situ. In addition, emerging in situ chromatin profiling can map eu-chromatin and hetero-chromatin states and reveal chromatin-state heterogeneity at cellular resolution, providing complementary information for refining regulatory annotations [97].

5.3. Regulatory Network Inference from Integrated Single-Cell Data

The integration of single-cell transcriptomics, chromatin accessibility profiling, and TF binding data provides a foundation for reconstructing GRNs underlying legume biology. When considering the nodulation process, trajectory inference and RNA velocity analyses have resolved two major developmental trajectories within *Medicago* nodules: a symbiotic lineage characterized by the stepwise activation of infection and nitrogen fixation genes, and a non-symbiotic branch diverging from distinct meristematic progenitor populations [59]. The bifurcation of these trajectories makes it possible to pinpoint transcriptional decisions that commit cortical cells either toward rhizobial accommodation or toward alternative developmental fates, offering a framework to study how regulatory inputs shape lineage divergence. To explore the regulatory logic underlying such divergent cell states, network modeling approaches, such as Single-Cell regulatory Network Inference and Clustering (SCENIC) and Multi-omics Integration for Network Inference with Accessibility and Chromatin data (MINI-AC) have been applied to predict TF–target interactions in a cell-type-specific manner [98,99]. For example, GRN inference consistently highlights *NIN* and *NF-YA* TFs as central hubs in infected cortical cells, confirming their established role in rhizobial entry and infection thread formation [12]. Beyond validating

known regulators, these GRNs predict additional TF modules associated with cortical cell proliferation, vascular differentiation, and senescence programs, thereby nominating new candidate regulators for experimental validation [2]. Patterns emerging from these predictions indicate that nodules rely on a layered regulatory architecture, in which infection-related TFs interface with modules controlling cell-cycle progression and vascular function, linking early infection cues to late-stage tissue differentiation. These predictions are strengthened by integrating GRN inference with chromatin accessibility and TF-binding datasets, enabling the construction of cell-type-resolved GRNs in legumes [4,100]. These networks reveal distinct yet interconnected transcriptional circuits across infection, vascular, and fixation zones, illustrating how spatially organized TF activities coordinate nodule development and symbiotic efficiency [14,59]. Such spatial partitioning of regulatory circuits demonstrates that nodules do not rely on a single master regulator but on distributed TF networks whose activity zones correspond to metabolic and developmental demands of each cell type. Collectively, these advances demonstrate that GRN inference provides not only confirmation of canonical symbiotic regulators but also a predictive framework to discover novel transcriptional players shaping legume cell identity and nitrogen-fixing capacity.

5.4. Perspectives and Future Directions

Recent advances in single-cell multi-omics have established foundational frameworks for dissecting cellular heterogeneity and regulatory mechanisms in legumes. Progress has been made in three major directions: (i) cell-type-specific expression profiling; (ii) mapping chromatin accessibility and TF–DNA interactions; (iii) inferring GRNs from integrated datasets. Studies on Lotus, Medicago, and soybean have generated high-resolution models that elucidate the developmental and symbiotic regulatory programs driving nodule formation and function [13,14,61]. Future research will benefit from combining spatial transcriptomics, three-dimensional chromatin conformation mapping, DNA methylation, and histone modification data in situ to capture higher-order genome architecture. In parallel, the development of cell-type-specific CRISPR/Cas9 genome editing systems will enable in planta validation of predicted regulatory modules [101]. The conceptual framework of the legume single-cell atlas synthesizes these functional genomic strategies with genetic variation, comparative cell atlases, and translational breeding applications, providing an integrative roadmap for future research and crop improvement.

6. Integrating eGWAS and $G \times E$ with Cell Atlases: Toward Predictive Breeding

Linking natural genetic variation to cell-type-specific expression traits is essential for translating single-cell discoveries into breeding applications. While single-cell atlases provide unprecedented cellular resolution, they currently lack integration with population-scale genomic datasets. Recent advances in single-cell genomics, combined with expression quantitative trait locus (eQTL) mapping and expression genome-wide association studies (eGWAS), will enable fine-mapping of regulatory variants underlying complex traits at single cell resolution. Integrating single-cell-resolution transcriptomic atlases with eGWAS and genotype–environment ($G \times E$) interaction analyses provide the opportunity to uncover the role of genes controlling a unique plant trait and expressed in specific cell types or that emerge only under particular environmental conditions, thereby refining our understanding of the genetic architecture of complex agronomic traits [102].

6.1. Cell-Type-Resolved eQTL and eGWAS Mapping

Combining GWAS with single-cell transcriptomic data enables a precise assessment of the effects of single-nucleotide polymorphisms (SNPs) on gene expression within defined cell types. In humans, the sc-eQTLGen consortium has pioneered efforts to resolve the impact of gene expression variation at the single-cell level on specific traits. For instance, using scRNA-seq of peripheral blood mononuclear cells (PBMCs), this consortium identified hundreds of cell-type-specific cis-eQTLs, including loci active exclusively in certain immune cell subsets, and reconstructed personalized gene regulatory networks across donors [103]. This work demonstrates the feasibility of connecting natural genetic variation to transcriptional regulation within defined cellular contexts. In plants, large-scale eQTL mapping has traditionally relied on bulk RNA-seq data. For example, Xu et al. [104] estimated cell-type proportions in Arabidopsis leaf tissues using computational deconvolution based on single-cell reference profiles, to identify over 15,038 cis-eQTLs, 62.83% of which showed significant cell-type interaction effects. Hence, the authors highlighted the importance of cell-type-aware modeling in capturing context-specific regulatory variation.

In legumes, cell-type-resolved eQTL datasets are missing. However, a recent preprint by Sun et al. [105] combined a soybean root scRNA-seq atlas with GWAS of root system architecture, revealing that gene networks predominantly expressed in the endodermis and metaploem were associated with root width and lateral root length. Earlier GWAS using the SoySNP50K panel also identified loci controlling root architecture [106]. The convergence between expression domains and trait-associated loci suggests that regulatory variants often act within narrowly defined cellular niches rather than at the whole-organ level, underscoring why bulk-based approaches fail to capture such effects. Together, these studies highlight the emerging potential of coupling high-density genotyping with single-cell datasets to accelerate causal discovery and guide precise breeding for enhanced nodulation, root architecture, and stress adaptation. As high-throughput single-cell genotyping and transcriptomics technologies become more accessible for crop species, we expect that cell-type-specific eQTL and eGWAS frameworks will soon enable fine-scale dissection of regulatory variants, bridging the gap between genotype, cell-type-specific expression patterns, and agronomic phenotypes in legumes. Together, these findings emphasize that integration of single-cell reference maps into population-scale genetic studies will uncover context-specific regulatory variants that cannot be captured by bulk approaches.

6.2. Single-Cell Resolution $G \times E$ Analysis

$G \times E$ interactions often emerge only in specific cell populations under stress, a complexity that bulk analyses cannot resolve. In Arabidopsis, cell-type-specific Ca^{2+} oscillations in the endodermis and pericycle under salt or osmotic stress highlight the heterogeneity of stress signaling [107]. More recently, multi-omics studies identified CBL-interacting protein kinase 23 (CIPK23) and NIN-Like Protein 7 (NLP7) as conserved drought response drivers across cell types [108]. The restricted activation of these pathways within cellular domains indicates that environmental responses are organized at the level of individual cell states rather than entire tissues, making single-cell approaches essential for attributing regulatory variation to its correct cellular context. Although single-cell $G \times E$ eQTL studies are lacking in legumes, these examples suggest that regulatory variation underlying nitrogen fixation and stress tolerance likely acts in a context-specific manner. Integrating legume single-cell atlases with GWAS/eGWAS across environments can help uncover stress-induced, cell-type-specific variants, advancing predictive breeding strategies.

6.3. Future Directions and Public Platforms

In the future, expanding the current whole-plant soybean single-cell transcriptome atlases to other legume species will be essential [2,4,5]. Complementary resources will also be required, such as multi-environment GWAS repositories, detailed phenotypic panels, and high-resolution genotype datasets. These resources will enable the integration of population genetics with cell-type-specific expression profiles. Equally important are computational platforms capable of handling large-scale, heterogeneous datasets. Cross-species integration tools (e.g., single-cell alignment mapping (SAMap), orthologous marker gene groups (OMGs), and species-agnostic transformers for universal cell embeddings and networks (SATURN)) need to be adapted for legumes to align cell states across diverse genetic backgrounds and environmental conditions. An integrated ecosystem of such datasets and analytical frameworks is especially critical for legumes, where species diversity, genome complexity, and trait-environment interactions often obscure causal relationships when analyses rely solely on bulk-level information. The establishment of open-access portals that host curated single-cell data, GWAS results, and trait-associated networks will allow researchers to perform reproducible and scalable analyses [109–111]. By unifying these resources, legume single-cell research can better resolve links between genetic variation and phenotypic outcomes at cellular resolution. This transition will enable the rational design of genome editing and breeding strategies targeting specific cell populations, ultimately improving nodulation efficiency, nutrient uptake, and stress resilience in legume crops.

7. Toward an Integrated Legume Single-Cell Biology: Cross-Species Integration and Synthetic Symbiosis

A comprehensive understanding of legume single-cell biology requires not only the accumulation of datasets but also the development of integrative frameworks that reveal conserved regulatory logic across species. The expansion of legume single-cell transcriptomic resources, combined with multi-omics profiling and synthetic biology, provides unique opportunities to dissect cell-type-specific GRNs. Yet, the field remains fragmented, with datasets produced under heterogeneous annotation standards that hinder cross-species comparisons. A unified framework, anchored in standardized cell ontologies, harmonized multi-species atlases, accurate metadata, and transferable synthetic regulatory modules, is essential to foster interoperability and accelerate the design of synthetic symbioses.

7.1. Standardizing Cell Ontologies and Annotation Systems

A major limitation of current legume single-cell studies is the absence of standardized cell ontologies. Datasets across species or even within a species often rely on study-specific clustering and marker genes, complicating integration and comparative analyses. Similar challenges in other fields have been addressed through ontology-based frameworks: the OBO Foundry has provided interoperable ontologies for biomedicine, and the Cell Ontology project has established hierarchical, reproducible classifications [112,113]. More recently, the minimum information required to annotate a cell type (MIRACL) standard introduced guidelines for reporting new cell types with essential metadata, enabling alignment with ontology entries [114]. A plant-specific equivalent of MIRACL, tailored to legumes, would enable consistent annotation, reproducibility, and cross-dataset integration. Such a framework would lay the foundation for multi-species legume cell atlases, advancing comparative studies of nodulation and symbiosis.

7.2. Cross-Species Atlas Integration and Conserved Regulatory Modules

Cross-species integration of single-cell datasets provides an approach to uncover evolutionarily conserved transcriptional programs and regulatory networks that con-

trol legume biology. While large-scale single-cell integration is well established in animal systems, where methods such as SAMap [111] align homologous cell types across distantly related species; similar efforts in plants are only beginning to emerge. Recent single-cell and single-nucleus atlases from legumes, including soybean roots and nodules [2,4] and Lotus [8], have identified conserved cell populations associated with rhizobial infection, nodule primordium initiation, and vascular remodeling. Consistently, Cervantes-Pérez et al. [58] transferred *Arabidopsis* markers to *Medicago* to resolve cell identities and early symbiotic responses, underscoring cross-species conservation at the cell-type level [58]. Comparative analyses demonstrate that key TFs, such as NIN and NLP family regulators, play conserved cell-type-specific expression across species. NIN is essential for nodulation even in the non-legume *Parasponia* [115], whereas NLPs mediate nitrate-dependent inhibition of nodulation in both Lotus and *Medicago* [7,116,117]. Together, these findings suggest the presence of a shared regulatory backbone governing nodulation. In addition to integration-based approaches, a recent framework named OMGs has been proposed to improve cross-species comparison of plant single-cell datasets [110]. By leveraging orthologous marker gene sets instead of relying solely on one-to-one ortholog alignment, OMGs robustly mapped homologous cell identities across 15 plant species, including multiple legumes, and revealed 14 broadly conserved cell-type groups spanning monocots and dicots. This method complements algorithms such as SAMap and provides a scalable solution for large-scale comparative plant single-cell transcriptomics, enabling the identification of conserved regulatory modules underlying developmental and symbiotic processes. Establishing robust pipelines for cross-species single-cell atlas integration by combining methods like SAMap and OMGs with homology-aware alignment algorithms such as SATURN [109] will be crucial for building a unified framework for legume cell identity annotation. These approaches will facilitate the discovery of deeply conserved regulatory circuits, illuminate lineage-specific innovations, and guide the selection or engineering of symbiotic gene modules for translational applications in crop breeding.

7.3. Synthetic Symbiosis and Transferable Trait Modules

Synthetic biology offers a compelling opportunity to reconstruct minimal symbiotic circuits and transfer modular regulatory programs from legumes into other plant species. Insights from single-cell atlases have enabled the identification of cell-type-specific regulatory modules that govern rhizobial infection, nodule initiation, and nitrogen-fixing activities [2,12,59,60]. These modules represent programmable elements that could drive symbiotic traits when expressed in heterologous systems. Foundational studies have shown that core nodulation regulators such as NIN and NF-YA TFs are capable, in principle, of triggering nodulation responses in non-legume models, although full functionality remains elusive [115]. In legumes, chromatin accessibility assays combined with transcriptomic analyses have begun to reveal cell-specific regulatory programs underlying symbiosis. For example, time-series (0–24 h) ATAC-seq and RNA-seq analyses in *Medicago* roots treated with rhizobial lipo-chitooligosaccharides (LCOs) were integrated with dynamic regulatory module networks to predict cis-regulatory elements and TFs, such as EIN3 and ERF1, whose RNAi knockdown reduces nodule numbers, supporting roles in early nodule development [100]. A particularly promising development comes from synthetic microbial-plant communication systems. Specifically, Boo et al. [118] engineered modular interkingdom communication channels in bacteria that sense environmental signals and transmit them to plants via small-molecule messengers, thereby activating plant gene expression. Although this study used *Arabidopsis* and *Solanum lycopersicum*, the modular signaling concept offers a scalable blueprint for encoding symbiotic triggers into microbial partners for legumes. Looking ahead, combining cell atlas-informed promoter modules with synthetic communi-

cation circuits and homology-aligned regulatory networks frameworks such as SAMap, OMGs and SATURN will enable assembly of transferable symbiotic trait modules. These could include multi-gene circuits for nodule initiation, rhizobial infection, and nutrient exchange, optimized to operate in specific cell types. Implemented through genome editing or transgenic expression in legumes or even cereals, such synthetic modules hold potential to reprogram plants for enhanced nitrogen fixation, improve stress resilience, and unlock symbiotic capabilities beyond traditional legume species. Incorporating spatial cell biology in situ helps integrate these single cell approaches by anchoring molecular profiles to their native tissue context. A conceptual framework summarizing these interconnected research directions is illustrated in Figure 3.

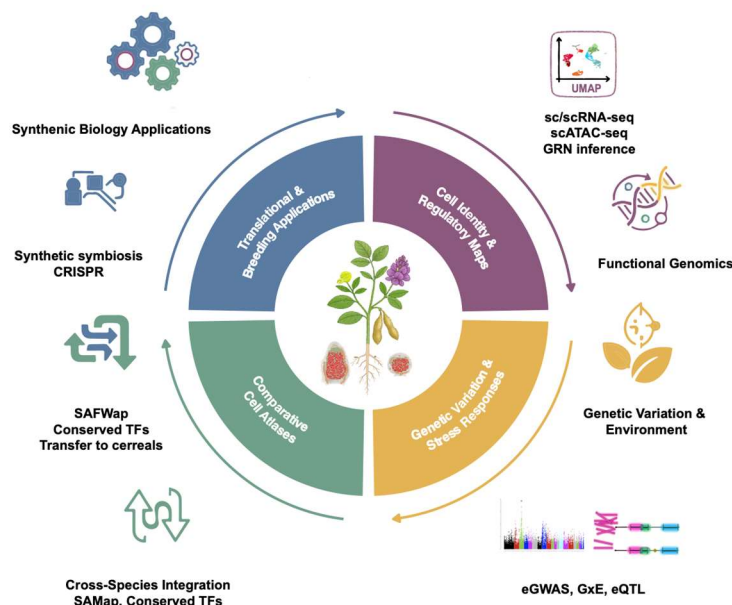


Figure 3. Conceptual framework of the Legume Single-Cell Atlas. The schem illustrates four interconnected research themes: (i) Cell Identity & Regulatory Maps (purple)—defined by scRNA-seq, snRNA-seq, ATAC-seq, and gene regulatory network inference; (ii) Genetic Variation & Stress Responses (yellow)—linking natural variation to cellular programs through eGWAS, G×E, and eQTL analyses; (iii) Comparative Cell Atlases (green)—cross-species integration using SAMap and conserved TFs; (iv) Translational & Breeding Applications (blue)—synthetic symbiosis, CRISPR-mediated engineering, and transfer of regulatory modules to cereals. Together, these directions highlight how single-cell genomics in legumes bridges basic regulatory mechanisms with translational applications in crop improvement.

Together, these advances mark a shift from descriptive single-cell atlases to the rational engineering of symbiotic traits. Integrating single-cell multi-omics with synthetic biology provides a roadmap toward programmable plant-microbe interactions and next-generation crop design.

8. Challenges and Future Directions

Despite remarkable progress in legume single-cell biology, several challenges continue to limit its broader application in plant science and crop improvement. Technical bottlenecks remain prominent: protoplast isolation and nuclei extraction often cause cell damage and low RNA yield, leading to sparse data that compromise the detection of low-abundance transcripts and the recovery of rare or transient states. Recent studies further highlight that even cells of the same type can adopt distinct transcriptional states under stress, for instance, immune-activated versus quiescent mesophyll cells during pathogen infection, or drought-induced mesophyll populations enriched either for ABA- or iron-responsive

programs. Such findings illustrate the importance of cell state transitions, yet systematic validation of state-specific regulators and their functional contribution to field-level traits is still lacking.

At the same time, current single-cell atlases remain largely descriptive because of the limited availability of robust cell-type markers, tissue-specific promoters, and high-throughput genome editing tools optimized for legumes. From a computational perspective, the integration of datasets across experiments, genotypes, and environmental conditions remains difficult due to variation in protocols, sequencing depth, and annotation standards. The lack of plant-specific ontologies and standardized metadata further constrains cross-species comparisons and the construction of unified multi-species legume atlases.

Addressing these limitations will require combined efforts. Integrative strategies that merge single-cell and spatial multi-omics with high-resolution mapping approaches (eQTL/eGWAS), advanced computational models of gene regulatory networks, and efficient genome editing platforms are needed to establish predictive pipelines. Such approaches will enable the identification and functional validation of key regulators acting at both the cell-type and cell-state levels. With these advances, single-cell research in legumes can move from descriptive atlases toward predictive and mechanistic models, creating opportunities for rational engineering of nodulation, nutrient uptake, and stress resilience, and ultimately contributing to sustainable crop improvement.

Author Contributions: Y.L. and M.S.H. contributed equally to the preparation of the manuscript. Y.L. and M.S.H. drafted the original manuscript. M.L. supervised the project and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the NSF awards #2414183 and #2425989 and the USDA-NIFA award #2022-67013-36144.

Data Availability Statement: All data are included in this article.

Acknowledgments: We are also grateful to Jared Vigil, Cam Chau Nguyen, and Rajashree Pradhan for their valuable comments on the manuscript. We apologize to those whose work was not cited because of space constraints.

Conflicts of Interest: M.L. is a consultant for INARI, a plant biotechnology company. The remaining authors declare no competing interests.

Abbreviations

The following abbreviations are used in this manuscript:

ABA	Absciscic acid
ACRs	Accessible chromatin regions
CSSP	Common Symbiotic Signaling Pathway
eGWAS	Expression genome-wide association studies
eQTL	Expression quantitative trait locus
FACS	Fluorescence-activated cell sorting
GRNs	Gene regulatory networks
GSTUs	Tau-class glutathione S-transferase
LCM	Laser capture microdissection
LCO	Lipo-chitooligosaccharides
MIRACL	Minimum information required to annotate a cell type
NF	Nod factor
NLR	Nucleotide-binding leucine-rich repeat
OMGs	Orthologous marker gene groups
PR	Pathogenesis-related
PRIMER	Pathogen-Responsive Immune Relay

PTI	Pattern-triggered immunity
QC	Quiescent center
RNS	Root nodule symbiosis
RNA-seq	RNA sequencing
ROS	Reactive oxygen species
SA	Salicylic acid
SAMap	Single-cell alignment mapping
SATURN	Species-agnostic transformers for universal cell embeddings and networks
sc/snRNA-seq	Single-cell/single-nucleus RNA sequencing
SMV	Soybean Mosaic Virus
snATAC-seq	Single-nucleus Assay for Transposase-Accessible Chromatin using sequencing
TFs	transcription factors

References

- Oldroyd, G.E.D.; Leyser, O. A Plant's Diet, Surviving in a Variable Nutrient Environment. *Science* **2020**, *368*, eaba0196. [\[CrossRef\]](#)
- Cervantes-Pérez, S.A.; Zogli, P.; Amini, S.; Thibivilliers, S.; Tennant, S.; Hossain, M.S.; Xu, H.; Meyer, I.; Nooka, A.; Ma, P.; et al. Single-Cell Transcriptome Atlases of Soybean Root and Mature Nodule Reveal New Regulatory Programs That Control the Nodulation Process. *Plant Commun.* **2024**, *5*, 100984. [\[CrossRef\]](#)
- Ye, K.; Bu, F.; Zhong, L.; Dong, Z.; Ma, Z.; Tang, Z.; Zhang, Y.; Yang, X.; Xu, X.; Wang, E.; et al. Mapping the Molecular Landscape of *Lotus japonicus* Nodule Organogenesis through Spatiotemporal Transcriptomics. *Nat. Commun.* **2024**, *15*, 6387. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhang, X.; Luo, Z.; Marand, A.P.; Yan, H.; Jang, H.; Bang, S.; Mendieta, J.P.; Minow, M.A.A.; Schmitz, R.J. A Spatially Resolved Multi-Omic Single-Cell Atlas of Soybean Development. *Cell* **2025**, *188*, 550–567.e19. [\[CrossRef\]](#) [\[PubMed\]](#)
- Fan, J.; Shen, Y.; Chen, C.; Chen, X.; Yang, X.; Liu, H.; Chen, R.; Liu, S.; Zhang, B.; Zhang, M.; et al. A Large-Scale Integrated Transcriptomic Atlas for Soybean Organ Development. *Mol. Plant* **2025**, *18*, 669–689. [\[CrossRef\]](#)
- Gonzalez-Rizzo, S.; Crespi, M.; Frugier, F. The *Medicago truncatula* CRE1 Cytokinin Receptor Regulates Lateral Root Development and Early Symbiotic Interaction with *Sinorhizobium meliloti*. *Plant Cell* **2006**, *18*, 2680–2693. [\[CrossRef\]](#)
- Lin, J.; Li, X.; Luo, Z.; Mysore, K.S.; Wen, J.; Xie, F. NIN Interacts with NLPs to Mediate Nitrate Inhibition of Nodulation in *Medicago truncatula*. *Nat. Plants* **2018**, *4*, 942–952. [\[CrossRef\]](#) [\[PubMed\]](#)
- Sun, Z.; Jiang, S.; Wang, D.; Li, L.; Liu, B.; Ran, Q.; Hu, L.; Xiong, J.; Tang, Y.; Gu, X.; et al. Single-Cell RNA-Seq of *Lotus japonicus* Provide Insights into Identification and Function of Root Cell Types of Legume. *J. Integr. Plant Biol.* **2023**, *65*, 1147–1152. [\[CrossRef\]](#)
- Cervantes-Pérez, S.A.; Thibivilliers, S.; Amini, S.; Pelletier, J.M.; Meyer, I.; Xu, H.; Tennant, S.; Ma, P.; Sprueill, C.M.; Farmer, A.D.; et al. Tabula Glycine: The Whole-Soybean Single-Cell Resolution Transcriptome Atlas. *bioRxiv* **2024**, 2024.07.08.602332. [\[CrossRef\]](#)
- Song, S.; Wang, J.; Zhou, J.; Cheng, X.; Hu, Y.; Wang, J.; Zou, J.; Zhao, Y.; Liu, C.; Hu, Z.; et al. Single-Cell RNA-Sequencing of Soybean Reveals Transcriptional Changes and Antiviral Functions of GmGSTU23 and GmGSTU24 in Response to Soybean Mosaic Virus. *Plant Cell Environ.* **2024**. online ahead of print. [\[CrossRef\]](#)
- Wang, L.; Zhou, Y.; Li, R.; Liang, J.; Tian, T.; Ji, J.; Chen, R.; Zhou, Y.; Fan, Q.; Ning, G.; et al. Single Cell-Type Transcriptome Profiling Reveals Genes That Promote Nitrogen Fixation in the Infected and Uninfected Cells of Legume Nodules. *Plant Biotechnol. J.* **2022**, *20*, 616–618. [\[CrossRef\]](#)
- Pereira, W.J.; Boyd, J.; Conde, D.; Triozzi, P.M.; Balmant, K.M.; Dervinis, C.; Schmidt, H.W.; Boaventura-Novaes, C.; Chakraborty, S.; Knaack, S.A.; et al. The Single-Cell Transcriptome Program of Nodule Development Cellular Lineages in *Medicago truncatula*. *Cell Rep.* **2024**, *43*, 113747. [\[CrossRef\]](#)
- Liu, Z.; Kong, X.; Long, Y.; Liu, S.; Zhang, H.; Jia, J.; Cui, W.; Zhang, Z.; Song, X.; Qiu, L.; et al. Integrated Single-Nucleus and Spatial Transcriptomics Captures Transitional States in Soybean Nodule Maturation. *Nat. Plants* **2023**, *9*, 515–524. [\[CrossRef\]](#)
- Frank, M.; Fechete, L.I.; Tedeschi, F.; Nadzieja, M.; Nørgaard, M.M.M.; Montiel, J.; Andersen, K.R.; Schierup, M.H.; Reid, D.; Andersen, S.U. Single-Cell Analysis Identifies Genes Facilitating Rhizobium Infection in *Lotus japonicus*. *Nat. Commun.* **2023**, *14*, 7171. [\[CrossRef\]](#)
- Guo, B.; Sun, L.; Jiang, S.; Ren, H.; Sun, R.; Wei, Z.; Hong, H.; Luan, X.; Wang, J.; Wang, X.; et al. Soybean Genetic Resources Contributing to Sustainable Protein Production. *Theor. Appl. Genet.* **2022**, *135*, 4095–4121. [\[CrossRef\]](#) [\[PubMed\]](#)
- Murphy, D.J. Agronomy and Environmental Sustainability of the Four Major Global Vegetable Oil Crops: Oil Palm, Soybean, Rapeseed, and Sunflower. *Agronomy* **2025**, *15*, 1465. [\[CrossRef\]](#)
- White, J.A.; Todd, J.; Newman, T.; Focks, N.; Girke, T.; de Ilárduya, O.M.; Jaworski, J.G.; Ohlrogge, J.B.; Benning, C. A New Set of Arabidopsis Expressed Sequence Tags from Developing Seeds. The Metabolic Pathway from Carbohydrates to Seed Oil. *Plant Physiol.* **2000**, *124*, 1582–1594. [\[CrossRef\]](#)

18. Benedito, V.A.; Torres-Jerez, I.; Murray, J.D.; Andrianakaja, A.; Allen, S.; Kakar, K.; Wandrey, M.; Verdier, J.; Zuber, H.; Ott, T.; et al. A Gene Expression Atlas of the Model Legume *Medicago truncatula*. *Plant J.* **2008**, *55*, 504–513. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Lister, R.; O'Malley, R.C.; Tonti-Filippini, J.; Gregory, B.D.; Berry, C.C.; Millar, A.H.; Ecker, J.R. Highly Integrated Single-Base Resolution Maps of the Epigenome in Arabidopsis. *Cell* **2008**, *133*, 523–536. [\[CrossRef\]](#)
20. Libault, M.; Farmer, A.; Joshi, T.; Takahashi, K.; Langley, R.J.; Franklin, L.D.; He, J.; Xu, D.; May, G.; Stacey, G. An Integrated Transcriptome Atlas of the Crop Model *Glycine max*, and Its Use in Comparative Analyses in Plants. *Plant J.* **2010**, *63*, 86–99. [\[CrossRef\]](#)
21. Severin, A.J.; Woody, J.L.; Bolon, Y.-T.; Joseph, B.; Diers, B.W.; Farmer, A.D.; Muehlbauer, G.J.; Nelson, R.T.; Grant, D.; Specht, J.E.; et al. RNA-Seq Atlas of *Glycine max*: A Guide to the Soybean Transcriptome. *BMC Plant Biol.* **2010**, *10*, 160. [\[CrossRef\]](#)
22. O'Rourke, J.A.; Iniguez, L.P.; Fu, F.; Bucciarelli, B.; Miller, S.S.; Jackson, S.A.; McClean, P.E.; Li, J.; Dai, X.; Zhao, P.X.; et al. An RNA-Seq Based Gene Expression Atlas of the Common Bean. *BMC Genom.* **2014**, *15*, 866. [\[CrossRef\]](#)
23. Afzal, M.; Alghamdi, S.S.; Migdadi, H.H.; Khan, M.A.; Nurmansyah; Mirza, S.B.; El-Harty, E. Legume Genomics and Transcriptomics: From Classic Breeding to Modern Technologies. *Saudi J. Biol. Sci.* **2020**, *27*, 543–555. [\[CrossRef\]](#)
24. Libault, M.; Farmer, A.; Brechenmacher, L.; Drnevich, J.; Langley, R.J.; Bilgin, D.D.; Radwan, O.; Neece, D.J.; Clough, S.J.; May, G.D.; et al. Complete Transcriptome of the Soybean Root Hair Cell, a Single-Cell Model, and Its Alteration in Response to *Bradyrhizobium japonicum* Infection. *Plant Physiol.* **2010**, *152*, 541–552. [\[CrossRef\]](#)
25. Belamkar, V.; Weeks, N.T.; Bharti, A.K.; Farmer, A.D.; Graham, M.A.; Cannon, S.B. Comprehensive Characterization and RNA-Seq Profiling of the HD-Zip Transcription Factor Family in Soybean (*Glycine max*) during Dehydration and Salt Stress. *BMC Genom.* **2014**, *15*, 950. [\[CrossRef\]](#)
26. Chen, W.; Yao, Q.; Patil, G.B.; Agarwal, G.; Deshmukh, R.K.; Lin, L.; Wang, B.; Wang, Y.; Prince, S.J.; Song, L.; et al. Identification and Comparative Analysis of Differential Gene Expression in Soybean Leaf Tissue under Drought and Flooding Stress Revealed by RNA-Seq. *Front. Plant Sci.* **2016**, *7*, 1044. [\[CrossRef\]](#)
27. Du, J.; Wang, S.; He, C.; Zhou, B.; Ruan, Y.-L.; Shou, H. Identification of Regulatory Networks and Hub Genes Controlling Soybean Seed Set and Size Using RNA Sequencing Analysis. *J. Exp. Bot.* **2017**, *68*, 1955–1972. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Goettel, W.; Xia, E.; Upchurch, R.; Wang, M.-L.; Chen, P.; An, Y.-Q.C. Identification and Characterization of Transcript Polymorphisms in Soybean Lines Varying in Oil Composition and Content. *BMC Genom.* **2014**, *15*, 299. [\[CrossRef\]](#)
29. Liu, A.; Xiao, Z.; Li, M.-W.; Wong, F.-L.; Yung, W.-S.; Ku, Y.-S.; Wang, Q.; Wang, X.; Xie, M.; Yim, A.K.-Y.; et al. Transcriptomic Reprogramming in Soybean Seedlings under Salt Stress. *Plant Cell Environ.* **2019**, *42*, 98–114. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Machado, F.B.; Moharana, K.C.; Almeida-Silva, F.; Gazara, R.K.; Pedrosa-Silva, F.; Coelho, F.S.; Grativol, C.; Venancio, T.M. Systematic Analysis of 1298 RNA-Seq Samples and Construction of a Comprehensive Soybean (*Glycine max*) Expression Atlas. *Plant J.* **2020**, *103*, 1894–1909. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Shen, Y.; Zhou, Z.; Wang, Z.; Li, W.; Fang, C.; Wu, M.; Ma, Y.; Liu, T.; Kong, L.-A.; Peng, D.-L.; et al. Global Dissection of Alternative Splicing in Paleopolyploid Soybean. *Plant Cell* **2014**, *26*, 996–1008. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Wang, L.; Liu, L.; Ma, Y.; Li, S.; Dong, S.; Zu, W. Transcriptome Profiling Analysis Characterized the Gene Expression Patterns Responded to Combined Drought and Heat Stresses in Soybean. *Comput. Biol. Chem.* **2018**, *77*, 413–429. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Xu, C.; Xia, C.; Xia, Z.; Zhou, X.; Huang, J.; Huang, Z.; Liu, Y.; Jiang, Y.; Casteel, S.; Zhang, C. Physiological and Transcriptomic Responses of Reproductive Stage Soybean to Drought Stress. *Plant Cell Rep.* **2018**, *37*, 1611–1624. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Khan, H.A.; Sharma, N.; Siddique, K.H.M.; Colmer, T.D.; Sutton, T.; Baumann, U. Comparative Transcriptome Analysis Reveals Molecular Regulation of Salt Tolerance in Two Contrasting Chickpea Genotypes. *Front. Plant Sci.* **2023**, *14*, 1191457. [\[CrossRef\]](#)
35. Kudapa, H.; Barmukh, R.; Garg, V.; Chitkineni, A.; Samineni, S.; Agarwal, G.; Varshney, R.K. Comprehensive Transcriptome Profiling Uncovers Molecular Mechanisms and Potential Candidate Genes Associated with Heat Stress Response in Chickpea. *Int. J. Mol. Sci.* **2023**, *24*, 1369. [\[CrossRef\]](#)
36. Malovichko, Y.V.; Shtark, O.Y.; Vasileva, E.N.; Nizhnikov, A.A.; Antonets, K.S. Transcriptomic Insights into Mechanisms of Early Seed Maturation in the Garden Pea (*Pisum sativum* L.). *Cells* **2020**, *9*, 779. [\[CrossRef\]](#)
37. Alves-Carvalho, S.; Aubert, G.; Carrère, S.; Cruaud, C.; Brochot, A.-L.; Jacquín, F.; Klein, A.; Martin, C.; Boucherot, K.; Kreplak, J.; et al. Full-Length de Novo Assembly of RNA-Seq Data in Pea (*Pisum sativum* L.) Provides a Gene Expression Atlas and Gives Insights into Root Nodulation in This Species. *Plant J.* **2015**, *84*, 1–19. [\[CrossRef\]](#)
38. Bhogireddy, S.; Xavier, A.; Garg, V.; Layland, N.; Arias, R.; Payton, P.; Nayak, S.N.; Pandey, M.K.; Puppala, N.; Varshney, R.K. Genome-Wide Transcriptome and Physiological Analyses Provide New Insights into Peanut Drought Response Mechanisms. *Sci. Rep.* **2020**, *10*, 4071. [\[CrossRef\]](#)
39. Wan, L.; Li, B.; Lei, Y.; Yan, L.; Ren, X.; Chen, Y.; Dai, X.; Jiang, H.; Zhang, J.; Guo, W.; et al. Mutant Transcriptome Sequencing Provides Insights into Pod Development in Peanut (*Arachis hypogaea* L.). *Front. Plant Sci.* **2017**, *8*, 1900. [\[CrossRef\]](#)
40. Rathod, V.; Hamid, R.; Tomar, R.S.; Patel, R.; Padhiyar, S.; Kheni, J.; Thirumalaisamy, P.P.; Munshi, N.S. Comparative RNA-Seq Profiling of a Resistant and Susceptible Peanut (*Arachis hypogaea*) Genotypes in Response to Leaf Rust Infection Caused by *Puccinia Arachidis*. *3 Biotech.* **2020**, *10*, 284. [\[CrossRef\]](#)

41. Zhang, H.; Wang, M.L.; Dang, P.; Jiang, T.; Zhao, S.; Lamb, M.; Chen, C. Identification of Potential QTLs and Genes Associated with Seed Composition Traits in Peanut (*Arachis hypogaea* L.) Using GWAS and RNA-Seq Analysis. *Gene* **2021**, *769*, 145215. [\[CrossRef\]](#)
42. Wang, L.; Rubio, M.C.; Xin, X.; Zhang, B.; Fan, Q.; Wang, Q.; Ning, G.; Becana, M.; Duanmu, D. CRISPR/Cas9 Knockout of Leghemoglobin Genes in *Lotus japonicus* Uncovers Their Synergistic Roles in Symbiotic Nitrogen Fixation. *New Phytol.* **2019**, *224*, 818–832. [\[CrossRef\]](#)
43. Handa, Y.; Nishide, H.; Takeda, N.; Suzuki, Y.; Kawaguchi, M.; Saito, K. RNA-Seq Transcriptional Profiling of an Arbuscular Mycorrhiza Provides Insights into Regulated and Coordinated Gene Expression in *Lotus japonicus* and *Rhizophagus irregularis*. *Plant Cell Physiol.* **2015**, *56*, 1490–1511. [\[CrossRef\]](#)
44. Luo, D.; Zhou, Q.; Wu, Y.; Chai, X.; Liu, W.; Wang, Y.; Yang, Q.; Wang, Z.; Liu, Z. Full-Length Transcript Sequencing and Comparative Transcriptomic Analysis to Evaluate the Contribution of Osmotic and Ionic Stress Components towards Salinity Tolerance in the Roots of Cultivated Alfalfa (*Medicago sativa* L.). *BMC Plant Biol.* **2019**, *19*, 32. [\[CrossRef\]](#)
45. Zhao, M.; Wang, T.; Sun, T.; Yu, X.; Tian, R.; Zhang, W.-H. Identification of Tissue-Specific and Cold-Responsive LncRNAs in *Medicago truncatula* by High-Throughput RNA Sequencing. *BMC Plant Biol.* **2020**, *20*, 99. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Jiang, X.; Yang, X.; Zhang, F.; Yang, T.; Yang, C.; He, F.; Gao, T.; Wang, C.; Yang, Q.; Wang, Z.; et al. Combining QTL Mapping and RNA-Seq Unravels Candidate Genes for Alfalfa (*Medicago sativa* L.) Leaf Development. *BMC Plant Biol.* **2022**, *22*, 485. [\[CrossRef\]](#)
47. Brown, A.V.; Conners, S.I.; Huang, W.; Wilkey, A.P.; Grant, D.; Weeks, N.T.; Cannon, S.B.; Graham, M.A.; Nelson, R.T. A New Decade and New Data at SoyBase, the USDA-ARS Soybean Genetics and Genomics Database. *Nucleic Acids Res.* **2021**, *49*, D1496–D1501. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Berendzen, J.; Brown, A.V.; Cameron, C.T.; Campbell, J.D.; Cleary, A.M.; Dash, S.; Hokin, S.; Huang, W.; Kalberer, S.R.; Nelson, R.T.; et al. The Legume Information System and Associated Online Genomic Resources. *Legume Sci.* **2021**, *3*, e74. [\[CrossRef\]](#)
49. Mun, T.; Bachmann, A.; Gupta, V.; Stougaard, J.; Andersen, S.U. Lotus Base: An Integrated Information Portal for the Model Legume *Lotus japonicus*. *Sci. Rep.* **2016**, *6*, 39447. [\[CrossRef\]](#)
50. Carrere, S.; Verdier, J.; Gamas, P. MtExpress, a Comprehensive and Curated RNAseq-Based Gene Expression Atlas for the Model Legume *Medicago truncatula*. *Plant Cell Physiol.* **2021**, *62*, 1494–1500. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Dash, S.; Cannon, E.K.S.; Kalberer, S.R.; Farmer, A.D.; Cannon, S.B. Chapter 8—PeanutBase and Other Bioinformatic Resources for Peanut. In *Peanuts*; Stalker, H.T., Wilson, R.F., Eds.; AOCS Press: Urbana, IL, USA, 2016; pp. 241–252. ISBN 978-1-63067-038-2.
52. Takanashi, K.; Takahashi, H.; Sakurai, N.; Sugiyama, A.; Suzuki, H.; Shibata, D.; Nakazono, M.; Yazaki, K. Tissue-Specific Transcriptome Analysis in Nodules of *Lotus japonicus*. *Mol. Plant Microbe Interact.* **2012**, *25*, 869–876. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Limpens, E.; Moling, S.; Hooiveld, G.; Pereira, P.A.; Bisseling, T.; Becker, J.D.; Küster, H. Cell- and Tissue-Specific Transcriptome Analyses of *Medicago truncatula* Root Nodules. *PLoS ONE* **2013**, *8*, e64377. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Roux, B.; Rodde, N.; Jardinaud, M.-F.; Timmers, T.; Sauviac, L.; Cottret, L.; Carrère, S.; Sallet, E.; Courcelle, E.; Moreau, S.; et al. An Integrated Analysis of Plant and Bacterial Gene Expression in Symbiotic Root Nodules Using Laser-Capture Microdissection Coupled to RNA Sequencing. *Plant J.* **2014**, *77*, 817–837. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Ryu, K.H.; Huang, L.; Kang, H.M.; Schiefelbein, J. Single-Cell RNA Sequencing Resolves Molecular Relationships Among Individual Plant Cells. *Plant Physiol.* **2019**, *179*, 1444–1456. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Shulze, C.N.; Cole, B.J.; Ciobanu, D.; Lin, J.; Yoshinaga, Y.; Gouran, M.; Turco, G.M.; Zhu, Y.; O'Malley, R.C.; Brady, S.M.; et al. High-Throughput Single-Cell Transcriptome Profiling of Plant Cell Types. *Cell Rep.* **2019**, *27*, 2241–2247.e4. [\[CrossRef\]](#)
57. Liu, H.; Hu, D.; Du, P.; Wang, L.; Liang, X.; Li, H.; Lu, Q.; Li, S.; Liu, H.; Chen, X.; et al. Single-Cell RNA-Seq Describes the Transcriptome Landscape and Identifies Critical Transcription Factors in the Leaf Blade of the Allotetraploid Peanut (*Arachis hypogaea* L.). *Plant Biotechnol. J.* **2021**, *19*, 2261–2276. [\[CrossRef\]](#)
58. Cervantes-Pérez, S.A.; Thibivilliers, S.; Laffont, C.; Farmer, A.D.; Frugier, F.; Libault, M. Cell-Specific Pathways Recruited for Symbiotic Nodulation in the *Medicago truncatula* Legume. *Mol. Plant* **2022**, *15*, 1868–1888. [\[CrossRef\]](#)
59. Ye, Q.; Zhu, F.; Sun, F.; Wang, T.-C.; Wu, J.; Liu, P.; Shen, C.; Dong, J.; Wang, T. Differentiation Trajectories and Biofunctions of Symbiotic and Un-Symbiotic Fate Cells in Root Nodules of *Medicago truncatula*. *Mol. Plant* **2022**, *15*, 1852–1867. [\[CrossRef\]](#)
60. Sun, B.; Wang, Y.; Yang, Q.; Gao, H.; Niu, H.; Li, Y.; Ma, Q.; Huan, Q.; Qian, W.; Ren, B. A High-Resolution Transcriptomic Atlas Depicting Nitrogen Fixation and Nodule Development in Soybean. *J. Integr. Plant Biol.* **2023**, *65*, 1536–1552. [\[CrossRef\]](#)
61. Liu, Z.; Yang, J.; Long, Y.; Zhang, C.; Wang, D.; Zhang, X.; Dong, W.; Zhao, L.; Liu, C.; Zhai, J.; et al. Single-Nucleus Transcriptomes Reveal Spatiotemporal Symbiotic Perception and Early Response in *Medicago*. *Nat. Plants* **2023**, *9*, 1734–1748. [\[CrossRef\]](#)
62. Cui, Y.; Su, Y.; Bian, J.; Han, X.; Guo, H.; Yang, Z.; Chen, Y.; Li, L.; Li, T.; Deng, X.W.; et al. Single-Nucleus RNA and ATAC Sequencing Analyses Provide Molecular Insights into Early Pod Development of Peanut Fruit. *Plant Commun.* **2024**, *5*, 100979. [\[CrossRef\]](#)
63. Deng, Q.; Du, P.; Gangurde, S.S.; Hong, Y.; Xiao, Y.; Hu, D.; Li, H.; Lu, Q.; Li, S.; Liu, H.; et al. scRNA-Seq Reveals Dark- and Light-Induced Differentially Expressed Gene Atlases of Seedling Leaves in *Arachis hypogaea* L. *Plant Biotechnol. J.* **2024**, *22*, 1848–1866. [\[CrossRef\]](#)

64. Liu, H.; Guo, Z.; Gangurde, S.S.; Garg, V.; Deng, Q.; Du, P.; Lu, Q.; Chitkineni, A.; Xiao, Y.; Wang, W.; et al. A Single-Nucleus Resolution Atlas of Transcriptome and Chromatin Accessibility for Peanut (*Arachis hypogaea* L.) Leaves. *Adv. Biol.* **2024**, *8*, 2300410. [\[CrossRef\]](#) [\[PubMed\]](#)
65. Wang, X.; Wang, R.; Huo, X.; Zhou, Y.; Umer, M.J.; Zheng, Z.; Jin, W.; Huang, L.; Li, H.; Yu, Q.; et al. Integration of Single-Nuclei Transcriptome and Bulk RNA-Seq to Unravel the Role of AhWRKY70 in Regulating Stem Cell Development in *Arachis hypogaea* L. *Plant Biotechnol. J.* **2025**, *23*, 1814–1831. [\[CrossRef\]](#)
66. Lepetit, M.; Brouquisse, R. Control of the Rhizobium–Legume Symbiosis by the Plant Nitrogen Demand Is Tightly Integrated at the Whole Plant Level and Requires Inter-Organ Systemic Signaling. *Front. Plant Sci.* **2023**, *14*, 1114840. [\[CrossRef\]](#)
67. Amini, S.; Doyle, J.J.; Libault, M. The Evolving Definition of Plant Cell Type. *Front. Plant Sci.* **2023**, *14*, 1271070. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Dorrity, M.W.; Alexandre, C.M.; Hamm, M.O.; Vigil, A.-L.; Fields, S.; Queitsch, C.; Cuperus, J.T. The Regulatory Landscape of *Arabidopsis thaliana* Roots at Single-Cell Resolution. *Nat. Commun.* **2021**, *12*, 3334. [\[CrossRef\]](#)
69. Marand, A.P.; Chen, Z.; Gallavotti, A.; Schmitz, R.J. A Cis-Regulatory Atlas in Maize at Single-Cell Resolution. *Cell* **2021**, *184*, 3041–3055.e21. [\[CrossRef\]](#)
70. Farmer, A.; Thibivilliers, S.; Ryu, K.H.; Schiefelbein, J.; Libault, M. Single-Nucleus RNA and ATAC Sequencing Reveals the Impact of Chromatin Accessibility on Gene Expression in *Arabidopsis* Roots at the Single-Cell Level. *Mol. Plant* **2021**, *14*, 372–383. [\[CrossRef\]](#)
71. Yin, R.; Xia, K.; Xu, X. Spatial Transcriptomics Drives a New Era in Plant Research. *Plant J.* **2023**, *116*, 1571–1581. [\[CrossRef\]](#)
72. Aida, M.; Beis, D.; Heidstra, R.; Willemsen, V.; Blilou, I.; Galinha, C.; Nussaume, L.; Noh, Y.-S.; Amasino, R.; Scheres, B. The *PLETHORA* Genes Mediate Patterning of the *Arabidopsis* Root Stem Cell Niche. *Cell* **2004**, *119*, 109–120. [\[CrossRef\]](#)
73. Cao, Y.; Ma, J.; Han, S.; Hou, M.; Wei, X.; Zhang, X.; Zhang, Z.J.; Sun, S.; Ku, L.; Tang, J.; et al. Single-Cell RNA Sequencing Profiles Reveal Cell Type-Specific Transcriptional Regulation Networks Conditioning Fungal Invasion in Maize Roots. *Plant Biotechnol. J.* **2023**, *21*, 1839–1859. [\[CrossRef\]](#)
74. Delannoy, E.; Batardiere, B.; Pateyron, S.; Soubigou-Taconnat, L.; Chiquet, J.; Colcombet, J.; Lang, J. Cell Specialization and Coordination in *Arabidopsis* Leaves upon Pathogenic Attack Revealed by ScRNA-Seq. *Plant Commun.* **2023**, *4*, 100676. [\[CrossRef\]](#)
75. Li, R.; Wang, Z.; Wang, J.-W.; Li, L. Combining Single-Cell RNA Sequencing with Spatial Transcriptome Analysis Reveals Dynamic Molecular Maps of Cambium Differentiation in the Primary and Secondary Growth of Trees. *Plant Commun.* **2023**, *4*, 100665. [\[CrossRef\]](#)
76. Nobori, T.; Oliva, M.; Lister, R.; Ecker, J.R. Multiplexed Single-Cell 3D Spatial Gene Expression Analysis in Plant Tissue Using PHYTOmap. *Nat. Plants* **2023**, *9*, 1026–1033. [\[CrossRef\]](#)
77. DeFalco, T.A.; Zipfel, C. Molecular Mechanisms of Early Plant Pattern-Triggered Immune Signaling. *Mol. Cell* **2021**, *81*, 3449–3467. [\[CrossRef\]](#)
78. Tang, B.; Feng, L.; Hulin, M.T.; Ding, P.; Ma, W. Cell-Type-Specific Responses to Fungal Infection in Plants Revealed by Single-Cell Transcriptomics. *Cell Host Microbe* **2023**, *31*, 1732–1747.e5. [\[CrossRef\]](#) [\[PubMed\]](#)
79. Abdelaal, T.; Mourragui, S.; Mahfouz, A.; Reinders, M.J.T. SpaGE: Spatial Gene Enhancement Using ScRNA-Seq. *Nucleic Acids Res.* **2020**, *48*, e107. [\[CrossRef\]](#) [\[PubMed\]](#)
80. Sun, Y.; Han, Y.; Sheng, K.; Yang, P.; Cao, Y.; Li, H.; Zhu, Q.-H.; Chen, J.; Zhu, S.; Zhao, T. Single-Cell Transcriptomic Analysis Reveals the Developmental Trajectory and Transcriptional Regulatory Networks of Pigment Glands in *Gossypium bickii*. *Mol. Plant* **2023**, *16*, 694–708. [\[CrossRef\]](#) [\[PubMed\]](#)
81. Cho, Y.; Kadam, U.; Park, B.; Amariillis, S.; Nguyen, K.-N.T.; Can, M.-H.T.; Lee, K.O.; Park, S.J.; Chung, W.S.; Hong, J.C. Recent Progress in Single-Cell Transcriptomic Studies in Plants. *Plant Biotechnol. Rep.* **2025**, *19*, 91–103. [\[CrossRef\]](#)
82. Nobori, T.; Monell, A.; Lee, T.A.; Sakata, Y.; Shirahama, S.; Zhou, J.; Nery, J.R.; Mine, A.; Ecker, J.R. A Rare PRIMER Cell State in Plant Immunity. *Nature* **2025**, *638*, 197–205. [\[CrossRef\]](#)
83. Yang, Q.; Li, Z.; Guan, K.; Wang, Z.; Tang, X.; Hong, Y.; Liu, Z.; Zhai, J.; Huang, A.; Long, Y.; et al. Comparative Single-Nucleus RNA-Seq Analysis Revealed Localized and Cell Type-Specific Pathways Governing Root-Microbiome Interactions. *Nat. Commun.* **2025**, *16*, 3169. [\[CrossRef\]](#)
84. Murray, J.D. Invasion by Invitation: Rhizobial Infection in Legumes. *Mol. Plant Microbe Interact.* **2011**, *24*, 631–639. [\[CrossRef\]](#)
85. Schiessl, K.; Lilley, J.L.S.; Lee, T.; Tamvakis, I.; Kohlen, W.; Bailey, P.C.; Thomas, A.; Luptak, J.; Ramakrishnan, K.; Carpenter, M.D.; et al. NODULE INCEPTION Recruits the Lateral Root Developmental Program for Symbiotic Nodule Organogenesis in *Medicago truncatula*. *Curr. Biol.* **2019**, *29*, 3657–3668.e5. [\[CrossRef\]](#)
86. Oldroyd, G.E.D.; Downie, J.A. Coordinating Nodule Morphogenesis with Rhizobial Infection in Legumes. *Annu. Rev. Plant Biol.* **2008**, *59*, 519–546. [\[CrossRef\]](#)
87. Roy, S.; Liu, W.; Nandety, R.S.; Crook, A.; Mysore, K.S.; Pislariu, C.I.; Frugoli, J.; Dickstein, R.; Udvardi, M.K. Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation. *Plant Cell* **2020**, *32*, 15–41. [\[CrossRef\]](#) [\[PubMed\]](#)

88. Serrano, K.; Bezruczyk, M.; Goudeau, D.; Dao, T.; O'Malley, R.; Malmstrom, R.R.; Visel, A.; Scheller, H.V.; Cole, B. Spatial Co-Transcriptomics Reveals Discrete Stages of the Arbuscular Mycorrhizal Symbiosis. *Nat. Plants* **2024**, *10*, 673–688. [\[CrossRef\]](#)
89. Díaz, V.; Villalobos, M.; Arriaza, K.; Flores, K.; Hernández-Saravia, L.P.; Velásquez, A. Decoding the Dialog Between Plants and Arbuscular Mycorrhizal Fungi: A Molecular Genetic Perspective. *Genes* **2025**, *16*, 143. [\[CrossRef\]](#) [\[PubMed\]](#)
90. Oldroyd, G.E.D. Speak, Friend, and Enter: Signalling Systems That Promote Beneficial Symbiotic Associations in Plants. *Nat. Rev. Microbiol.* **2013**, *11*, 252–263. [\[CrossRef\]](#) [\[PubMed\]](#)
91. Jean-Baptiste, K.; McFaline-Figueroa, J.L.; Alexandre, C.M.; Dorrity, M.W.; Saunders, L.; Bubba, K.L.; Trapnell, C.; Fields, S.; Queitsch, C.; Cuperus, J.T. Dynamics of Gene Expression in Single Root Cells of *Arabidopsis thaliana*. *Plant Cell* **2019**, *31*, 993–1011. [\[CrossRef\]](#)
92. Soyano, T.; Shimoda, Y.; Kawaguchi, M.; Hayashi, M. A Shared Gene Drives Lateral Root Development and Root Nodule Symbiosis Pathways in Lotus. *Science* **2019**, *366*, 1021–1023. [\[CrossRef\]](#)
93. Wijman, T.; Mohd-Radzman, N.A.; Liu, J.; Oldroyd, G.E.D.; Kohlen, W.; Kulikova, O.; Heidstra, R.; Scheres, B.; Franssen, H.J. The *Medicago truncatula* lncRNA ENOD40 Is a Mediator of MicroRNA169-Controlled NF-YA Activity in Nodule Initiation. *Proc. Natl. Acad. Sci. USA* **2025**, *122*, e2424231122. [\[CrossRef\]](#)
94. Fu, M.; Sun, J.; Li, X.; Guan, Y.; Xie, F. Asymmetric Redundancy of Soybean Nodule Inception (NIN) Genes in Root Nodule Symbiosis. *Plant Physiol.* **2022**, *188*, 477–489. [\[CrossRef\]](#)
95. Li, X.; Zhou, H.; Cheng, L.; Ma, N.; Cui, B.; Wang, W.; Zhong, Y.; Liao, H. Shoot-to-Root Translocated GmNN1/FT2a Triggers Nodulation and Regulates Soybean Nitrogen Nutrition. *PLoS Biol.* **2022**, *20*, e3001739. [\[CrossRef\]](#) [\[PubMed\]](#)
96. Jiao, W.; Wang, M.; Guan, Y.; Guo, W.; Zhang, C.; Wei, Y.; Zhao, Z.; Ma, H.; Wang, L.; Jiang, X.; et al. Transcriptional Regulatory Network Reveals Key Transcription Factors for Regulating Agronomic Traits in Soybean. *Genome Biol.* **2024**, *25*, 313. [\[CrossRef\]](#)
97. Caballero, L.; Pasternak, T.; Riyazuddin, R.; Pérez-Pérez, J.M. Connecting High-Resolution 3D Chromatin Maps with Cell Division and Cell Differentiation at the Root Apical Meristem. *Plant Cell Rep.* **2024**, *43*, 232. [\[CrossRef\]](#) [\[PubMed\]](#)
98. Aibar, S.; González-Blas, C.B.; Moerman, T.; Huynh-Thu, V.A.; Imrichova, H.; Hulselmans, G.; Rambow, F.; Marine, J.-C.; Geurts, P.; Aerts, J.; et al. SCENIC: Single-Cell Regulatory Network Inference and Clustering. *Nat. Methods* **2017**, *14*, 1083–1086. [\[CrossRef\]](#) [\[PubMed\]](#)
99. Pérez, N.M.; Ferrari, C.; Engelhorn, J.; Depuydt, T.; Nelissen, H.; Hartwig, T.; Vandepoele, K. MINI-AC: Inference of Plant Gene Regulatory Networks Using Bulk or Single-Cell Accessible Chromatin Profiles. *Plant J.* **2024**, *117*, 280–301. [\[CrossRef\]](#)
100. Knaack, S.A.; Conde, D.; Chakraborty, S.; Balmant, K.M.; Irving, T.B.; Maia, L.G.S.; Triozzi, P.M.; Dervinis, C.; Pereira, W.J.; Maeda, J.; et al. Temporal Change in Chromatin Accessibility Predicts Regulators of Nodulation in *Medicago truncatula*. *BMC Biol.* **2022**, *20*, 252. [\[CrossRef\]](#)
101. Gao, J.-P.; Su, Y.; Jiang, S.; Liang, W.; Lou, Z.; Frugier, F.; Xu, P.; Murray, J.D. Applying Conventional and Cell-Type-Specific CRISPR/Cas9 Genome Editing in Legume Plants. *ABIOTECH* **2025**, *6*, 346–360. [\[CrossRef\]](#)
102. Kremling, K.A.G.; Chen, S.-Y.; Su, M.-H.; Lepak, N.K.; Romay, M.C.; Swarts, K.L.; Lu, F.; Lorient, A.; Bradbury, P.J.; Buckler, E.S. Dysregulation of Expression Correlates with Rare-Allele Burden and Fitness Loss in Maize. *Nature* **2018**, *555*, 520–523. [\[CrossRef\]](#)
103. van der Wijst, M.G.P.; de Vries, D.H.; Groot, H.E.; Trynka, G.; Hon, C.C.; Bonder, M.J.; Stegle, O.; Nawijn, M.C.; Idaghdour, Y.; van der Harst, P.; et al. The Single-Cell EQTLGen Consortium. *Elife* **2020**, *9*, e52155. [\[CrossRef\]](#)
104. Xu, C.; Song, L.-Y.; Zhou, Y.; Ma, D.-N.; Ding, Q.-S.; Guo, Z.-J.; Li, J.; Song, S.-W.; Zhang, L.-D.; Zheng, H.-L. Integration of EQTL and GWAS Analysis Uncovers a Genetic Regulation of Natural Ionomic Variation in Arabidopsis. *Plant Cell Rep.* **2023**, *42*, 1473–1485. [\[CrossRef\]](#)
105. Sun, Y.; Miller, C.; Rajurkar, A.B.; Lynch, R.C.; Alyward, A.; Zhang, L.; Shaner, M.; Copeland, C.D.; Ye, H.; Nguyen, H.T.; et al. Genome-Wide Association Study Reveals Influence of Cell-Specific Gene Networks on Soybean Root System Architecture. *bioRxiv* **2024**, 2024.02.27.581071. [\[CrossRef\]](#)
106. Seck, W.; Torkamaneh, D.; Belzile, F. Comprehensive Genome-Wide Association Analysis Reveals the Genetic Basis of Root System Architecture in Soybean. *Front. Plant Sci.* **2020**, *11*, 590740. [\[CrossRef\]](#) [\[PubMed\]](#)
107. Kiegle, E.; Moore, C.A.; Haseloff, J.; Tester, M.A.; Knight, M.R. Cell-Type-Specific Calcium Responses to Drought, Salt and Cold in the Arabidopsis Root. *Plant J.* **2000**, *23*, 267–278. [\[CrossRef\]](#) [\[PubMed\]](#)
108. Liu, M.; Xu, Y.; Song, Y.; Fan, D.; Li, J.; Zhang, Z.; Wang, L.; He, J.; Chen, C.; Ma, C. Hierarchical Regulatory Networks Reveal Conserved Drivers of Plant Drought Response at the Cell-Type Level. *Adv. Sci.* **2025**, *12*, 2415106. [\[CrossRef\]](#)
109. Rosen, Y.; Brbić, M.; Roohani, Y.; Swanson, K.; Li, Z.; Leskovec, J. Toward Universal Cell Embeddings: Integrating Single-Cell RNA-Seq Datasets across Species with SATURN. *Nat. Methods* **2024**, *21*, 1492–1500. [\[CrossRef\]](#)
110. Chau, T.N.; Timilsena, P.R.; Bathala, S.P.; Kundu, S.; Bargmann, B.O.R.; Li, S. Orthologous Marker Groups Reveal Broad Cell Identity Conservation across Plant Single-Cell Transcriptomes. *Nat. Commun.* **2025**, *16*, 201. [\[CrossRef\]](#)
111. Tarashansky, A.J.; Musser, J.M.; Khariton, M.; Li, P.; Arendt, D.; Quake, S.R.; Wang, B. Mapping Single-Cell Atlases throughout Metazoa Unravels Cell Type Evolution. *Elife* **2021**, *10*, e66747. [\[CrossRef\]](#)

112. Smith, B.; Ashburner, M.; Rosse, C.; Bard, J.; Bug, W.; Ceusters, W.; Goldberg, L.J.; Eilbeck, K.; Ireland, A.; Mungall, C.J.; et al. The OBO Foundry: Coordinated Evolution of Ontologies to Support Biomedical Data Integration. *Nat. Biotechnol.* **2007**, *25*, 1251–1255. [[CrossRef](#)] [[PubMed](#)]
113. Diehl, A.D.; Meehan, T.F.; Bradford, Y.M.; Brush, M.H.; Dahdul, W.M.; Dougall, D.S.; He, Y.; Osumi-Sutherland, D.; Ruttenberg, A.; Sarntinijai, S.; et al. The Cell Ontology 2016: Enhanced Content, Modularization, and Ontology Interoperability. *J. Biomed. Semant.* **2016**, *7*, 44. [[CrossRef](#)] [[PubMed](#)]
114. Lubiana, T.; Roncaglia, P.; Mungall, C.J.; Quardokus, E.M.; Fortriede, J.D.; Osumi-Sutherland, D.; Diehl, A.D. Guidelines for Reporting Cell Types: The MIRACL Standard. *arXiv* **2022**, arXiv:2204.09673. [[CrossRef](#)]
115. Bu, F.; Rutten, L.; Roswanjaya, Y.P.; Kulikova, O.; Rodriguez-Franco, M.; Ott, T.; Bisseling, T.; van Zeijl, A.; Geurts, R. Mutant Analysis in the Nonlegume *Parasponia andersonii* Identifies NIN and NF-YA1 Transcription Factors as a Core Genetic Network in Nitrogen-Fixing Nodule Symbioses. *New Phytol.* **2020**, *226*, 541–554. [[CrossRef](#)]
116. Nishida, H.; Tanaka, S.; Handa, Y.; Ito, M.; Sakamoto, Y.; Matsunaga, S.; Betsuyaku, S.; Miura, K.; Soyano, T.; Kawaguchi, M.; et al. A NIN-LIKE PROTEIN Mediates Nitrate-Induced Control of Root Nodule Symbiosis in *Lotus japonicus*. *Nat. Commun.* **2018**, *9*, 499. [[CrossRef](#)]
117. Misawa, F.; Ito, M.; Nosaki, S.; Nishida, H.; Watanabe, M.; Suzuki, T.; Miura, K.; Kawaguchi, M.; Suzaki, T. Nitrate Transport via NRT2.1 Mediates NIN-LIKE PROTEIN-Dependent Suppression of Root Nodulation in *Lotus japonicus*. *Plant Cell* **2022**, *34*, 1844–1862. [[CrossRef](#)]
118. Boo, A.; Toth, T.; Yu, Q.; Pfothenhauer, A.; Fields, B.D.; Lenaghan, S.C.; Stewart, C.N.; Voigt, C.A. Synthetic Microbe-to-Plant Communication Channels. *Nat. Commun.* **2024**, *15*, 1817. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.