



# Single-cell RNA-seq of *Lotus japonicus* provide insights into identification and function of root cell types of legume

Zhanmin Sun<sup>1†</sup>, Sanjie Jiang<sup>2†</sup>, Dan Wang<sup>3†</sup>, Linxia Li<sup>1</sup>, Boxin Liu<sup>2</sup>, Qifan Ran<sup>4</sup>, Lizhen Hu<sup>5</sup>, Junbo Xiong<sup>6</sup>, Yixiong Tang<sup>1</sup>, Xiaofeng Gu<sup>1</sup>, Yanmin Wu<sup>1\*</sup> and Zhe Liang<sup>1\*</sup>

1. Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, China

2. BGI Genomics, Shenzhen 518083, China

3. Key Laboratory of Southwest China Wildlife Resources Conservation, China West Normal University, Nanchong 637002, China

4. Chongqing Academy of Animal Sciences, Chongqing 402460, China

5. Institute of Animal and Veterinary Science, Jiangxi Academy of Agricultural Sciences, Nanchang 330200, China

6. Hubei Key Laboratory of Animal Embryo and Molecular Breeding, Institute of Animal and Veterinary Science, Hubei Academy of Agricultural Science, Wuhan 430072, China

†Zhanmin Sun, Sanjie Jiang, and Dan Wang contributed equally to this work.

\*Correspondences: Yanmin Wu ([wuyanmin@caas.cn](mailto:wuyanmin@caas.cn)); Zhe Liang ([liangzhe@caas.cn](mailto:liangzhe@caas.cn)). Dr. Liang is fully responsible for the distribution of the materials associated with this article)



Zhanmin Sun



Zhe Liang

## ABSTRACT

The roots of legume plant play a crucial role in nitrogen fixation. However, the transcriptomes of different cell types of legume root and their functions remain largely unknown. Here, we performed single-cell RNA sequencing and profiled

more than 22,000 single cells from root tips of *Lotus japonicus*, a model species of legume. We identified seven clusters corresponding to seven major cell types, which were validated by *in situ* hybridization. Further analysis revealed regulatory programs including phytohormone and nodulation associated with specific cell types, and revealed conserved and diverged features for the cell types. Our results represent the first single-cell resolution transcriptome for legume root tips and a valuable resource for studying the developmental and physiological functions of various cell types in legumes.

Keywords: genome, legume, *Lotus japonicus*, single-cell RNA-seq

Sun, Z., Jiang, S., Wang, D., Li, L., Liu, B., Ran, Q., Hu, L., Xiong, J., Tang, Y., Gu, X., Wu, Y., and Liang, Z. (2023). Single-cell RNA-seq of *Lotus japonicus* provide insights into identification and function of root cell types of legume. *J. Integr. Plant Biol.* **65**: 1147–1152.

## INTRODUCTION

Single-cell RNA sequencing (scRNA-seq) is a powerful tool for analyzing unique gene expression patterns to understand cellular heterogeneity and for exploring the functions of cell types at single-cell resolution (Klein et al., 2015; Zeisel et al., 2015; Tirosh et al., 2016). Recently, scRNA-seq has been applied to several plant species and revealed divergent features between the monocots and dicots root tip transcriptomes (Zhang et al., 2019; Liu et al., 2021), emphasizing

the importance of analyzing scRNA-seq across diverse model plant species. The roots of legume plants host symbiotic bacteria and have important roles in nitrogen fixation. However, the molecular basis of legume root cell types and their functions are largely unknown. Therefore, we applied scRNA-seq to root tips of *Lotus japonicus*, an important perennial model legume for studying symbiotic nitrogen fixation (Andersen and Stougaard, 2022), to decipher the function of various cell types of roots, for understanding root development and nodulation at single-cell resolution.

## RESULTS AND DISCUSSION

First, we improved the *L. japonicus* genome assembly and annotation compared to the previous version (Li et al., 2020). We sequenced *L. japonicus* genome using Illumina short-read sequencing, PacBio long-read sequencing, 10x Genomics, and chromatin conformation capture (Hi-C) sequencing. The 540.04-Mb genome was assembled into six chromosomes and consisting of 37 scaffolds with an N50 of 4.43 Mb (Figure S1; Tables S1–S10). We also sequenced micro RNAs (miRNAs) and long non-coding RNAs (lncRNAs), respectively, and sequenced messenger RNAs (mRNAs) from different tissues using a combination of Illumina and PacBio sequencing, and predicted 38,684 high-confidence putative protein-coding genes and 8,312 non-coding RNAs (ncRNA) (Figure S1; Tables S11–S18).

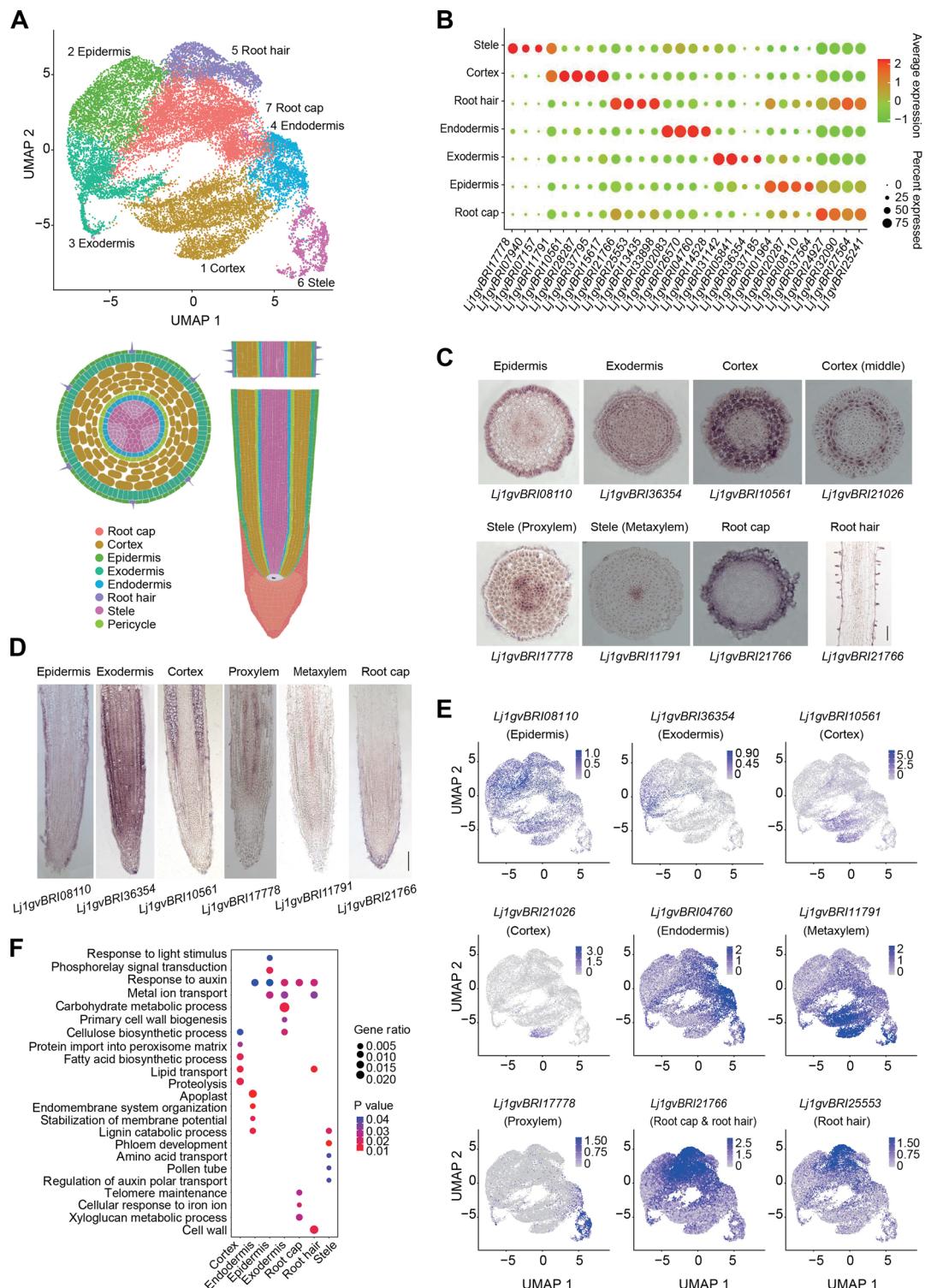
Based on our high-quality assembly and annotation of the *L. japonicus* genome, we used scRNA-seq to analyze gene expression in individual *L. japonicus* root tip cells (Figure S2A). Protoplasts were processed using the 10x Genomics platform (Figure S2B, C). We captured 22,688 root cells with 3,183 (median number) genes per cell (Table S19), and found that the global gene expression profiles of combined scRNA-seq and bulk RNA-seq from independently isolated root tip cells or protoplasts are highly correlated ( $r > 0.8$ ; Figure S2D–G). After non-linear dimensional reduction by the Uniform Manifold Approximation and Projection (UMAP) and *t*-distributed stochastic neighbor embedding algorithm, and further unsupervised clustering analyses, the dataset grouped root cells into seven distinct clusters (Figures 1A, S3A), each containing between 1,126 and 6,760 cells with different cell cycle statuses (Table S20). Seven similar clusters were also observed when we removed 4,074 differentially expressed genes (DEGs) in response to protoplasting (Table S21, Figure S3B, C), suggesting that the effect of protoplasting to clustering is minor. We identified a series of marker genes that are specifically expressed in one or two clusters (Figure 1B; Table S22). To annotate these clusters, we performed RNA *in situ* hybridizations and identified bona fide cell-type marker genes for most of *L. japonicus* root major cell types (Figure 1C, D). For example, *Lj1gvBRI08110* (encoding a 14-3-3-like protein) was specifically expressed in the cells of epidermis, *Lj1gvBRI36354* (encoding a polygalacturonase) in exodermis cells, *Lj1gvBRI10561* (encoding a lipid transfer protein) in cortex cells, *Lj1gvBRI21766* (encoding a peroxidase) in root cap and root hair, *Lj1gvBRI17778* (encoding a peroxidase) in protoxylem, and *Lj1gvBRI11791* (encoding an unknown protein) in metaxylem (Figure 1E). Notably, cell-type clusters 1 and 6 contain root cap and root hair cells. Considering that root hair marker genes are highly conserved between monocots and dicots (Liu et al., 2021), we used orthologs of conserved *Arabidopsis* and rice root hair marker genes *Lj1gvBRI25553*, *Lj1gvBRI33898*, and *Lj1gvBRI13435*, to annotate cluster 6 as root hair cells and

cluster 1 as root cap cells. The remaining cell-type cluster 5 was further annotated as endodermis cells using orthologs of *Arabidopsis* marker genes *Lj1gvBRI04760*, *Lj1gvBRI26570*, and *Lj1gvBRI22927* (Figures 1A, B, E, S3A, B, D, E). Next, we annotated a subcluster within the cortex that contains many cells by identifying a novel marker gene, *Lj1gvBRI21026* (encoding a eukaryotic aspartyl protease), which is specifically expressed in the middle cortex (second and third cell layers of cortex) (Figures 1C, S5A). In summary, our scRNA-seq dataset revealed seven distinct clusters corresponding to seven major cell types, which were validated by *in situ* hybridization experiments.

To provide insights into the functions of each cell type in *L. japonicus* root, we performed Gene Ontology (GO) analysis and found matching functional categories for each cell type (Figures 1F, S4, S5B), which further validated our cell-type cluster annotation. For example, genes expressed in the cortex implicated in processes related to “lipid transport”, “proteolysis”, and “fatty acid biosynthetic process”. Genes involved in “phloem development” were enriched in the stele, while genes related to “endomembrane system organization” were enriched in endodermis (Figures 1F, S4). On the other hand, genes that were expressed in all clusters include functions related mainly to “intracellular protein transport” and “signal peptide processing” (Figure S4).

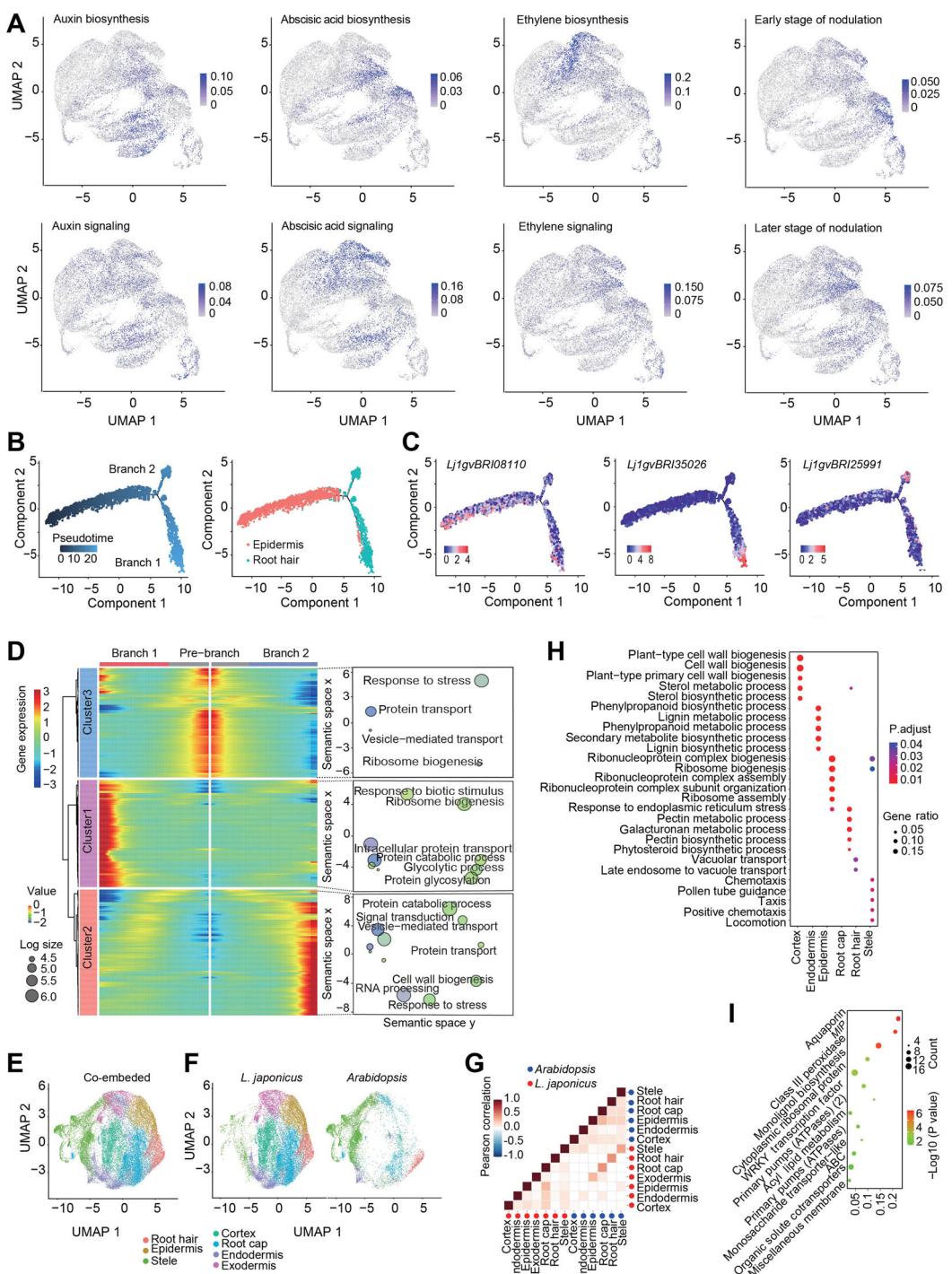
Phytohormones play important roles in modulating root nodule symbiosis (Liu et al., 2018; Lin et al., 2020). We analyzed the spatiotemporal expression patterns of genes involved in phytohormone and nodulation functions (Table S23), and found that genes related to auxin, abscisic acid, cytokinin, gibberellin (GA), and brassinosteroid (BR) biosynthesis were enriched in root cap, cortex, and endodermis, whereas those of genes related to ethylene biosynthesis were enriched in root cap and root hair. Most genes potentially involved in plant hormone signaling were broadly distributed in all clusters, except for those related to jasmonic acid signaling, which were enriched in root cap, endodermis, and cortex (Figures 2A, S3F). Root nodulation is a key process in legume species. At the early nodulation stage (Roy et al., 2020), including processes of root hair deformation and membrane depolarization, genes involved in symbiotic nitrogen fixation (SNF) were enriched in root hair, endodermis, and stele. At the late nodulation stage, including processes of nodule development and maturation, SNF genes mostly were enriched in root cap and cortex cells (Figure 2A; Table S24), suggesting distinct functions of SNF genes during nodule development. These cell-type specific expression patterns provide new insights into the functions of these genes, and suggest diverse phytohormone and nodulation functions that could be taking place in different cell types of *L. japonicus* root tips.

Our trajectory analysis demonstrated a gradual transition in the gene expression during the differentiation from



**Figure 1.** Single-cell RNA sequencing and cluster annotation of *Lotus japonicus* root tips

**Figure 1. Single-cell RNA sequencing and cluster annotation of *Lotus japonicus* root tips**  
**(A)** Uniform Manifold Approximation and Projection (UMAP) visualization of putative clusters from 22,688 cells (top) and the spatial distribution of cell types (bottom) in root tips of *L. japonicus*. Each dot indicates a single cell. Colors in the root tip diagram indicate corresponding cell clusters. **(B)** Expression of cell-type marker genes for each cluster. Dot diameter, proportion of cluster cells expressing a given gene; color, mean expression across cells in that cluster. **(C, D)** Cross section **(C)** and longitudinal section **(D)** for representative RNA *in situ* hybridization of cell-type marker genes for the seven single cell-type clusters. Scale bars, 50 µm. **(E)** UMAP visualization of cell-type marker genes for the seven cell-type clusters in *L. japonicus*. Color bars indicate scaled expression levels. **(F)** Scatter plots showing Gene Ontology term enrichment analysis of differentially expressed genes in each cell type.

**Figure 2. Function and evolutionary analysis of *Lotus japonicus* single cells**

**(A)** Uniform Manifold Approximation and Projection (UMAP) visualization of expression patterns of the transcripts of genes related to auxin, abscisic acid, ethylene, cytokinin, and nodulation functions, respectively. Color bars indicate scaled expression levels. **(B)** Differentiation trajectory of epidermal and root hair cells. Colors of the dots indicate the pseudotime score. **(C)** Expression of cell-type marker genes on the pseudotime trajectory. Color bars indicate scaled expression levels. **(D)** Heatmap showing the expression of the 605 most significant differentially expressed genes (DEGs) across the pseudotime trajectory. Each row represents one gene and the DEGs are clustered into three clusters with distinct expression patterns. Color bars indicate the relative expression levels. Cluster 1 genes are enriched in cells on Branch 1 of the trajectory, Cluster 2 genes are enriched in cells on Branch 2 of the trajectory, and Cluster 3 genes are enriched in cells on the Prebranch. Scatter plots of representative Gene Ontology (GO) terms for each cluster are shown on the right. **(E and F)** UMAP visualization of *L. japonicus* and *Arabidopsis* clusters (E, co-embedded, F, separated) after alignment. The colors indicate cell types. **(G)** Heatmap showing Pearson's correlations between *L. japonicus* and *Arabidopsis* cell-type transcriptomes. The single-cell data were merged together for each cell type prior to comparison. **(H)** Scatter plots of GO term enrichment analysis of conserved expressed genes for each cell type in *Arabidopsis* and *L. japonicus*. The color of the q value represents the significance of enrichment, and the size of the filled circles is proportional to gene number. **(I)** Gene families with the most divergent expression patterns. Divergent genes are defined as gene expression differences with >5-fold change. Hypergeometric distribution test was used as a significance enrichment test.

epidermis to *L. japonicus* root hair cells (Figure 2B,C). We identified 604 DEGs across the pseudotime order. These genes fell into three clusters with distinct gene expression patterns reflecting transcriptional rewiring during root hair development (Figure 2D). The Prebranch in this trajectory included epidermis cells and genes related to “response to stress” and “vesicle-mediated transport.” Branch 1 consists of root hair cells, and genes were related to response to “biotic stimulus” and “protein transport” (Figure 2D). Branch 2 included root hair cells and genes related to “small GTPase mediated signal transduction.” The divergence of root hair cells into two branches suggested we should explore the differences in gene expression in root hair cells between branches. Notably, we found that Branch 2 was significantly enriched in genes involved in nodulation ( $P < 0.01$ , hypergeometric test), such as *Lj1gvBRI17393* which encodes a nodule pectate lyase required for the degradation of plant cell walls during rhizobial infection (Xie et al., 2012); and *Lj1gvBRI05019* which encodes a phytochrome that modulates nitric oxide concentration during symbiosis (Fukudome et al., 2016), and *Lj1gvBRI14528* which encodes an isoflavone synthase homologous to *GmIFS* that is required for nodulation in soybean (Subramanian et al., 2006); and *Lj1gvBRI23668* which encodes a molybdenum transporter homologous to *Medicago truncatula MtMOT1.3* which mediates molybdate uptake into infected cells and is essential for SNF (Tejada-Jimenez et al., 2017). However, these genes were not detected in Branch 1 and the Prebranch, and UMAP showed that cells from Branch 1 and Branch 2 were separated into two subclusters of root hair cells (Figure S5C, D). These results suggested divergent functions among root hair cells, and the nodulation genes preferentially expressed in Branch 2 cells, consistent with a previous study reporting nodulation activity in some root hair cells but not others (Heidstra et al., 1994; Pawlowski and Demchenko, 2012).

To examine the evolutionary conservation of root cell-type development, we compared scRNA-seq datasets between *L. japonicus* and *Arabidopsis* (Zhang et al., 2019) by aligning the orthologous genes between them, and then clustering their root cell types. Consistent cell-type homologies were identified based on shared cluster membership. UMAP visualization and correlation analyses revealed that clusters of root hair, stele, and root cap cells are more similar than other cell types between *L. japonicus* and *Arabidopsis*, which may help to predict properties and conservation of homologous cell types (Figure 2E, F). Notably, the orthologs of marker genes of *L. japonicus* such as root hair cell-specific *Lj1gvBRI11214* and *Lj1gvBRI22367*, stele cell-specific *Lj1gvBRI02219* and *Lj1gvBRI12470*, and root cap cell-specific *Lj1gvBRI05693* and *Lj1gvBRI05344*, were also specifically expressed in corresponding cell types in *Arabidopsis* (Figure S6A), suggesting their evolutionarily conserved functions in root hair, root cap, and stele across species (Figure 2G). Although we previously showed that the majority of root tip cell types were not conserved between rice and *Arabidopsis* (Liu et al., 2021), our new

result suggests the functions expressed in cell types are more conserved within dicots than between monocots and dicots. Next, we analyzed divergent and conserved genes for each cell type between *L. japonicus* and *Arabidopsis* (Figure 2H, I). The most divergent gene families included those encoding members of the aquaporin family, the major intrinsic protein (MIP) family, and class III peroxidases (Figure 2I). Notably, the same most-diverged gene families were also observed between rice and *Arabidopsis* (Liu et al., 2021), suggesting that the diverged gene families noted in pairwise comparisons might have contributed to the divergence of cell functions and root morphology in different plant lineages.

In summary, our single-cell transcriptomic analysis has identified robust cell-type markers and cell-type specific regulatory programs for most of the cell types we identified in *L. japonicus* root tip. The evolutionary conserved and diverged cell-type and species-specific features of gene expression between *L. japonicus* and *Arabidopsis* will provide new insights into the functions and development of cell types in *L. japonicus*. Our study presents the transcriptomic landscape of major cell types of legume root tips at single-cell resolution and a valuable resource to study function and evolution of cell types in legume.

#### Data availability statement

All sequencing data used in this study are available on Sequence Read Archive (SRA) through accession numbers SRP376527, SRP376541, SRP376509, SRP377130, SRP377126, SRP377348, SRP377535, SRP377357, SRP377359, SRP377372, SRP377335, CNA0050696.

## ACKNOWLEDGEMENTS

We thank Yongzhen Pang for sharing *L. japonicus* seeds. We thank Chenxi Zhao for helping in assembly and annotation of the genome of *L. japonicus*. This work was supported by a survey on Central China grassland forage germplasm resources (2017FY100604), National Natural Science Foundation of China (32001395), Chongqing Natural Science Foundation (stc2020jcyjmsxmX0626), The Agricultural Science and Technology Innovation Program (ASTIP) (to Z.L.).

## CONFLICTS OF INTEREST

The authors declare no conflict of interest.

## AUTHOR CONTRIBUTIONS

Z.S., Y.W., and Z.L. conceived and designed the research. Z.S. and D.W. conducted experiments. Z.S., S.J., L.L., B.L., Q.R., L.H., J.X., Y.X., X.G., Y.W. and Z.L. analyzed the data. Z.S., Y.W. and Z.L. wrote the paper. All authors read and approved of the manuscript.

**Edited by:** Jianbing Yan, Huazhong Agricultural University, China

**Received** Oct. 3, 2022; **Accepted** Dec. 16, 2022; **Published** Dec. 20, 2022

## REFERENCES

- Andersen, S.U., and Stougaard, J. (2022). *Lotus japonicus*. Curr. Biol. **32**: R149–R150.
- Fukudome, M., Calvo-Begueria, L., Kado, T., Osuki, K., Rubio, M.C., Murakami, E., Nagata, M., Kucho, K., Sandal, N., Stougaard, J., Becana, M., and Uchiumi, T. (2016). Hemoglobin LjGlb1-1 is involved in nodulation and regulates the level of nitric oxide in the *lotus japonicus*-*mesorhizobium loti* symbiosis. J. Exp. Bot. **67**: 5275–5283.
- Heidstra, R., Geurts, R., Franssen, H., Spaink, H.P., Van Kammen, A., and Bisseling, T. (1994). Root hair deformation activity of nodulation factors and their fate on *vicia sativa*. Plant Physiol. **105**: 787–797.
- Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., and Kirschner, M.W. (2015). Droplet bar-coding for single-cell transcriptomics applied to embryonic stem cells. Cell **161**: 1187–1201.
- Li, H., Jiang, F., Wu, P., Wang, K., and Cao, Y. (2020). A high-quality genome sequence of model legume *Lotus japonicus* (MG-20) provides insights into the evolution of root nodule symbiosis. Genes **11**: 483.
- Lin, J., Frank, M., and Reid, D. (2020). No home without hormones: How plant hormones control legume nodule organogenesis. Plant Commun. **1**: 100104.
- Liu, H., Zhang, C., Yang, J., Yu, N., and Wang, E. (2018). Hormone modulation of legume-rhizobial symbiosis. J. Integr. Plant Biol. **60**: 632–648.
- Liu, Q., Liang, Z., Feng, D., Jiang, S.J., Wang, Y.F., Du, Z.Y., Li, R.X., Hu, G.H., Zhang, P.X., Ma, Y.F., Lohmann, J.U., and Gu, X.F. (2021). Transcriptional landscape of rice roots at the single-cell resolution. Mol. Plant **14**: 384–394.
- Pawlowski, K., and Demchenko, K.N. (2012). The diversity of actinorhizal symbiosis. Protoplasma **249**: 967–979.
- Roy, S., Liu, W., Nandety, R.S., Crook, A., Mysore, K.S., Pislaru, C.I., Frugoli, J., Dickstein, R., and Udvardi, M.K. (2020). Celebrating 20 years of genetic discoveries in legume nodulation and symbiotic nitrogen fixation. Plant Cell **32**: 15–41.
- Subramanian, S., Stacey, G., and Yu, O. (2006). Endogenous isoflavones are essential for the establishment of symbiosis between soybean and *Bradyrhizobium japonicum*. Plant J. **48**: 261–273.
- Tejada-Jimenez, M., Gil-Diez, P., Leon-Mediavilla, J., Wen, J., Mysore, K.S., Imperial, J., and Gonzalez-Guerrero, M. (2017). *Medicago truncatula* molybdate transporter type 1 (MtMOT1.3) is a plasma membrane molybdenum transporter required for nitrogenase activity in root nodules under molybdenum deficiency. New Phytol. **216**: 1223–1235.
- Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., 2nd, Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., Fallahi-Sichani, M., Dutton-Regester, K., Lin, J.R., Cohen, O., Shah, P., Lu, D., Genshaft, A.S., Hughes, T.K., Ziegler, C.G., Kazer, S.W., Gaillard, A., Kolb, K.E., Villani, A.C., Johannessen, C.M., Andreev, A.Y., Van Allen, E.M., Bertagnolli, M., Sorger, P.K., Sullivan, R.J., Flaherty, K.T., Frederick, D.T., Jane-Valbuena, J., Yoon, C.H., Rozenblatt-Rosen, O., Shalek, A.K., Regev, A., and Garraway, L.A. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science **352**: 189–196.

Xie, F., Murray, J.D., Kim, J., Heckmann, A.B., Edwards, A., Oldroyd, G.E., and Downie, J.A. (2012). Legume pectate lyase required for root infection by *rhizobia*. Proc. Natl. Acad. Sci. U.S.A. **109**: 633–638.

Zeisel, A., Munoz-Manchado, A.B., Codeluppi, S., Lonnerberg, P., La Manno, G., Jureus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., Rolny, C., Castelo-Branco, G., Hjerling-Leffler, J., and Linnarsson, S. (2015). Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science **347**: 1138–1142.

Zhang, T.Q., Xu, Z.G., Shang, G.D., Wang, J.W. (2019). A single-cell RNA sequencing profiles the developmental landscape of *Arabidopsis* root. Mol. plant **12**: 648–660.

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article: <http://onlinelibrary.wiley.com/doi/10.1111/jipb.13435/supinfo>

**Figure S1.** *Lotus japonicus* genome assembly and annotation

**Figure S2.** Morphology, single-cell RNA sequencing (scRNA-seq), and bulk RNA-seq of *Lotus japonicus* root tip

**Figure S3.** Single-cell RNA sequencing (scRNA-seq) cluster and annotation

**Figure S4.** Gene Ontology term analysis of differentially expressed genes (DEGs) in each cell type and commonly expressed genes in several cell types of *Lotus japonicus* root tip

**Figure S5.** Analysis of the subclusters of root cortex and root hair cells

**Figure S6.** Analysis of conserved cell types and marker genes between *Lotus japonicus* and *Arabidopsis*. Statistics of repeat sequences in *L. japonicus* genome

**Table S1.** Statistics of Illumina RNA sequencing raw reads

**Table S2.** Statistics of Illumina RNA sequencing clean reads

**Table S3.** Statistics of full-length isoform sequencing (Iso-seq) of *Lotus japonicus*

**Table S4.** Statistics of 10x Genomics sequencing of *Lotus japonicus*

**Table S5.** Statistics of PacBio sequencing of *Lotus japonicus*

**Table S6.** Statistics of chromatin conformation capture (Hi-C) sequencing of *Lotus japonicus*

**Table S7.** Statistics of the assembled contigs and scaffolds of *Lotus japonicus* (10x Genomics)

**Table S8.** The mapping statistics of the high-quality chromatin conformation capture (Hi-C) data

**Table S9.** Completeness of the genome based on BUSCO (Benchmarking Universal Single-Copy Orthologs) analysis

**Table S10.** The summary of genome assembly results of *Lotus japonicus*

**Table S11.** Summary of annotated genes

**Table S12.** Summary of functional annotation of predicted genes

**Table S13.** Summary of annotated non-coding RNA

**Table S14.** List of annotated long non-coding (lncRNA) in *Lotus japonicus*

**Table S15.** List of annotated small nuclear RNA (snRNA) in *Lotus japonicus*

**Table S16.** List of annotated ribosomal RNA (rRNA) in *Lotus japonicus*

**Table S17.** List of annotated transfer RNA (tRNA) in *Lotus japonicus*

**Table S18.** Statistics of repeat sequences in *Lotus japonicus* genome

**Table S19.** Statistics of single-cell RNA sequencing (scRNA-seq) in *Lotus japonicus*

**Table S20.** Number of cells captured by single-cell RNA sequencing (scRNA-seq) for each cell type

**Table S21.** Differentially expressed genes (DEGs) between protoplasted bulk RNA sequencing (RNA-seq) and un-protoplasted (root tissues) bulk RNA-seq

**Table S22.** List of enriched genes for each cluster of *Lotus japonicus*

**Table S23.** List of phytohormone-related genes

**Table S24.** List of nodulation-related genes

**Table S25.** List of primers used in this study