Supplementary Files

Title:

Single-cell RNA-seq Describes the Transcriptome Landscape and Identifies Critical Transcription Factors in the Leaf Blade of the Allotetraploid Peanut (*Arachis hypogaea L.*)

Authors Name:

Hao Liu[1, 4], Dongxiu Hu[1, 4], Puxuan Du[1, 4], Liping Wang[1], Xuanqiang Liang[1], Haifen Li[1], Qing Lu[1], Shaoxiong Li[1], Haiyan Liu[1], Xiaoping Chen[1*], Rajeev K Varshney[2,3,*], Yanbin Hong[1*]

*Corresponding Authors

Affiliations:

[1]Guangdong Provincial Key Laboratory of Crop Genetic Improvement, South China Peanut Sub-Center of National Center of Oilseed Crops Improvement, Crops Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou, Guangdong Province, 510640 China.

[2]Center of Excellence in Genomics & Systems Biology, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, Telangana, 502324 India.

[3]State Agricultural Biotechnology Centre, Centre for Crop and Food Innovation, Food Futures Institute, Murdoch University, Murdoch, Western Australia, 6150 Australia.

[4] These authors contributed equally to this article.

Running title

The Development and Application of Single-cell RNA-seq on Peanut Leaf Cells.

Supplementary Figures

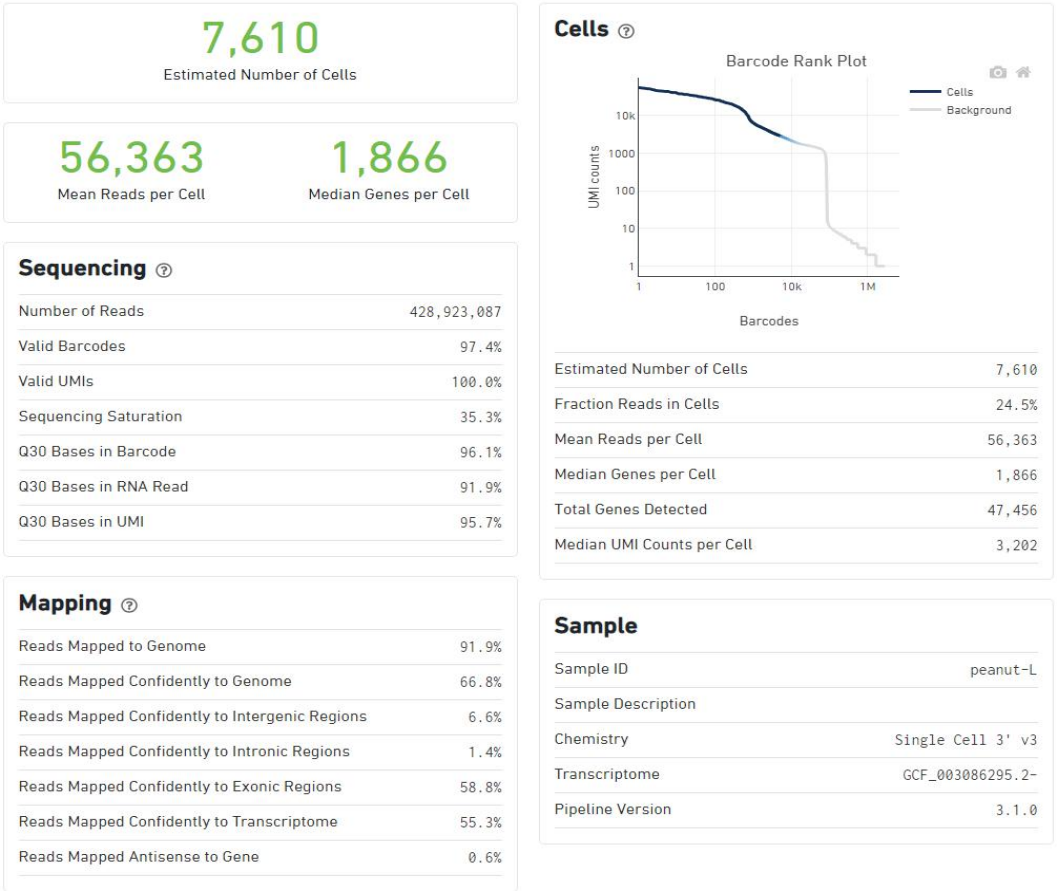

Figure S1. Brief chart of Cell Ranger software report.

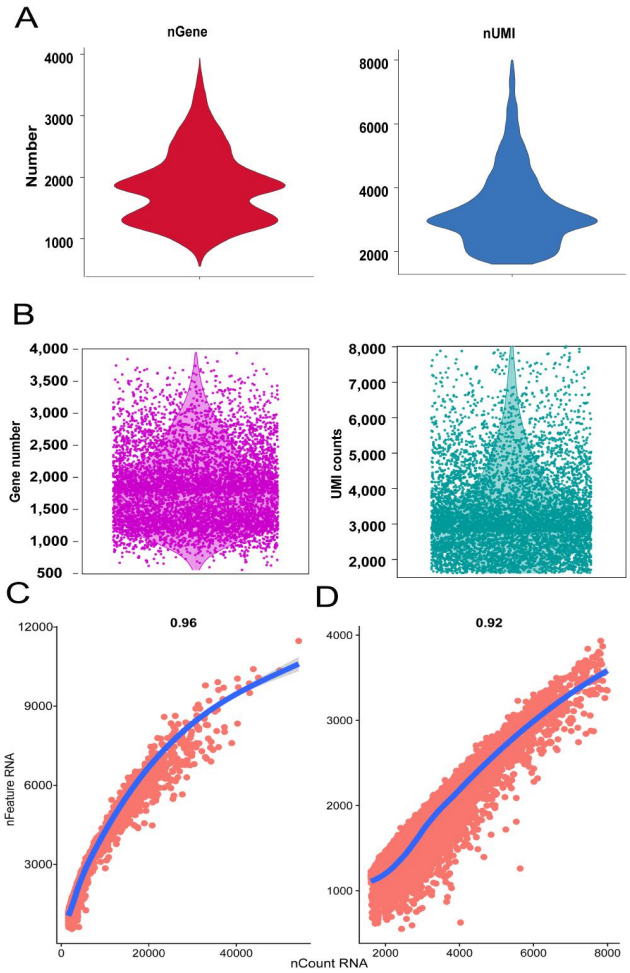

Figure S2. Data quality control determined the medium number of gene and UMI. A,violin plots. B, dot plots. C, Correlation coefficient between the UMI and gene, before filter. D, Correlation coefficient of the UMI and gene, after filter.

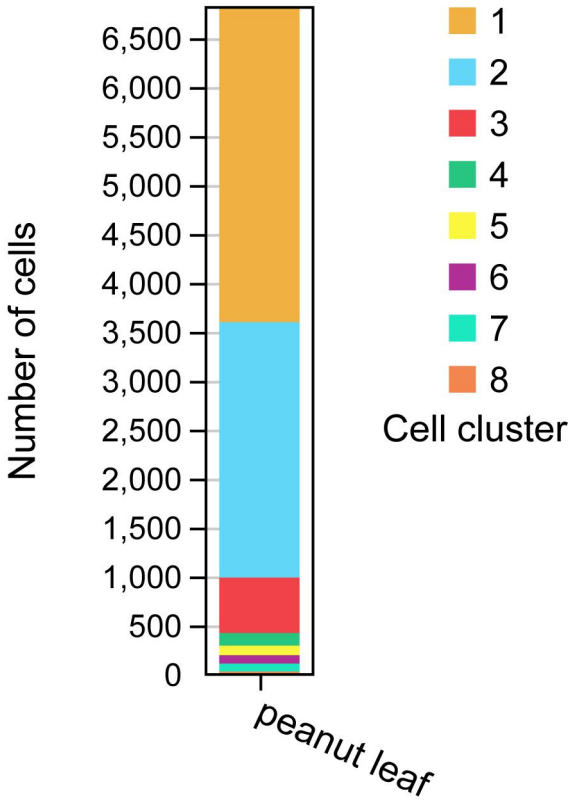

Figure S3. The number of cells in each peanut leaf cell cluster.

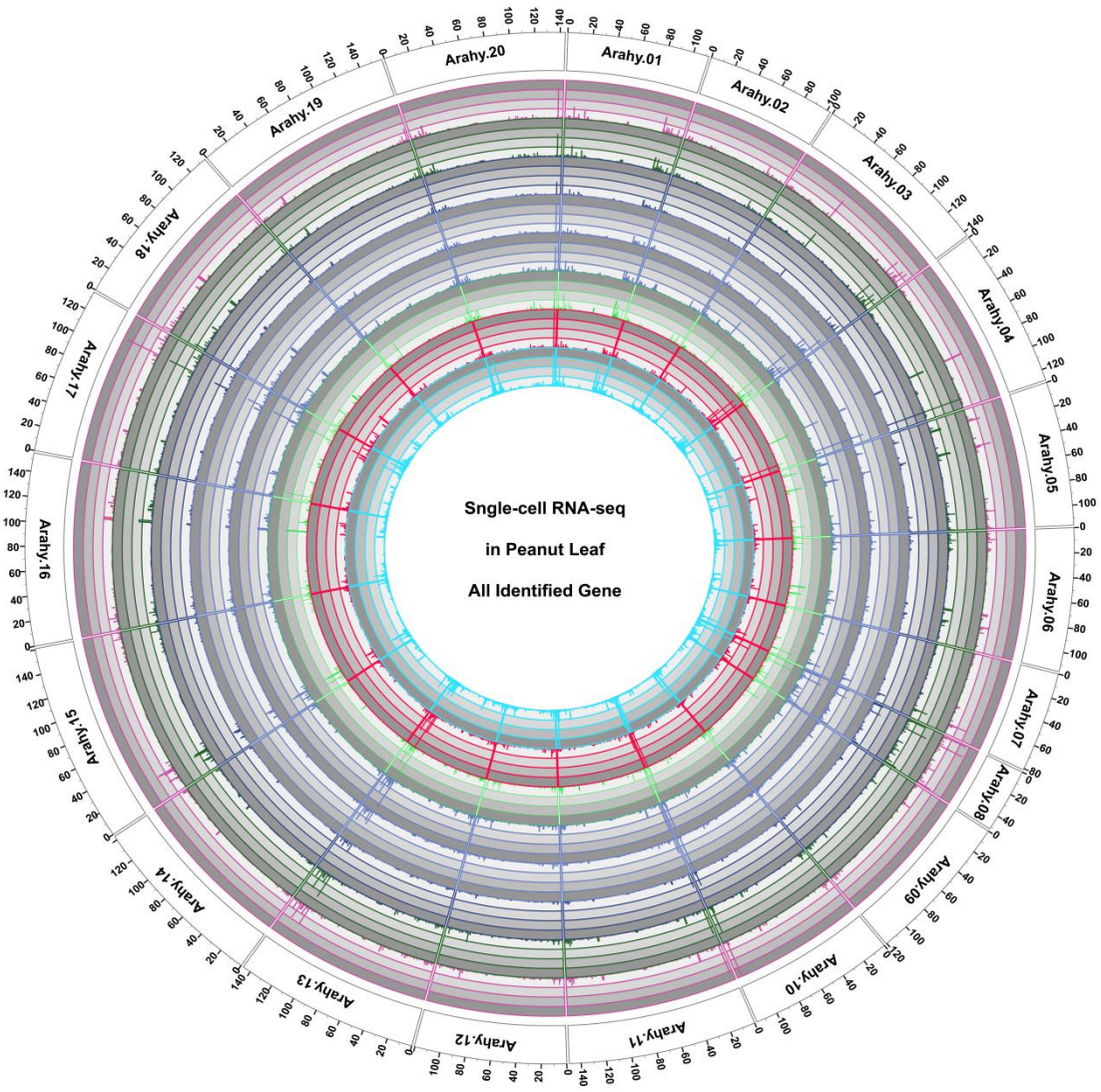

Figure S4. Circular visualization displaying the global transcriptional analysis of all identified genes (relative expression level) in peanut leaf cell by scRNA-seq (**Table S3**). Form outer circular to inner circular represented the cell cluster 1 to 8, respectively.

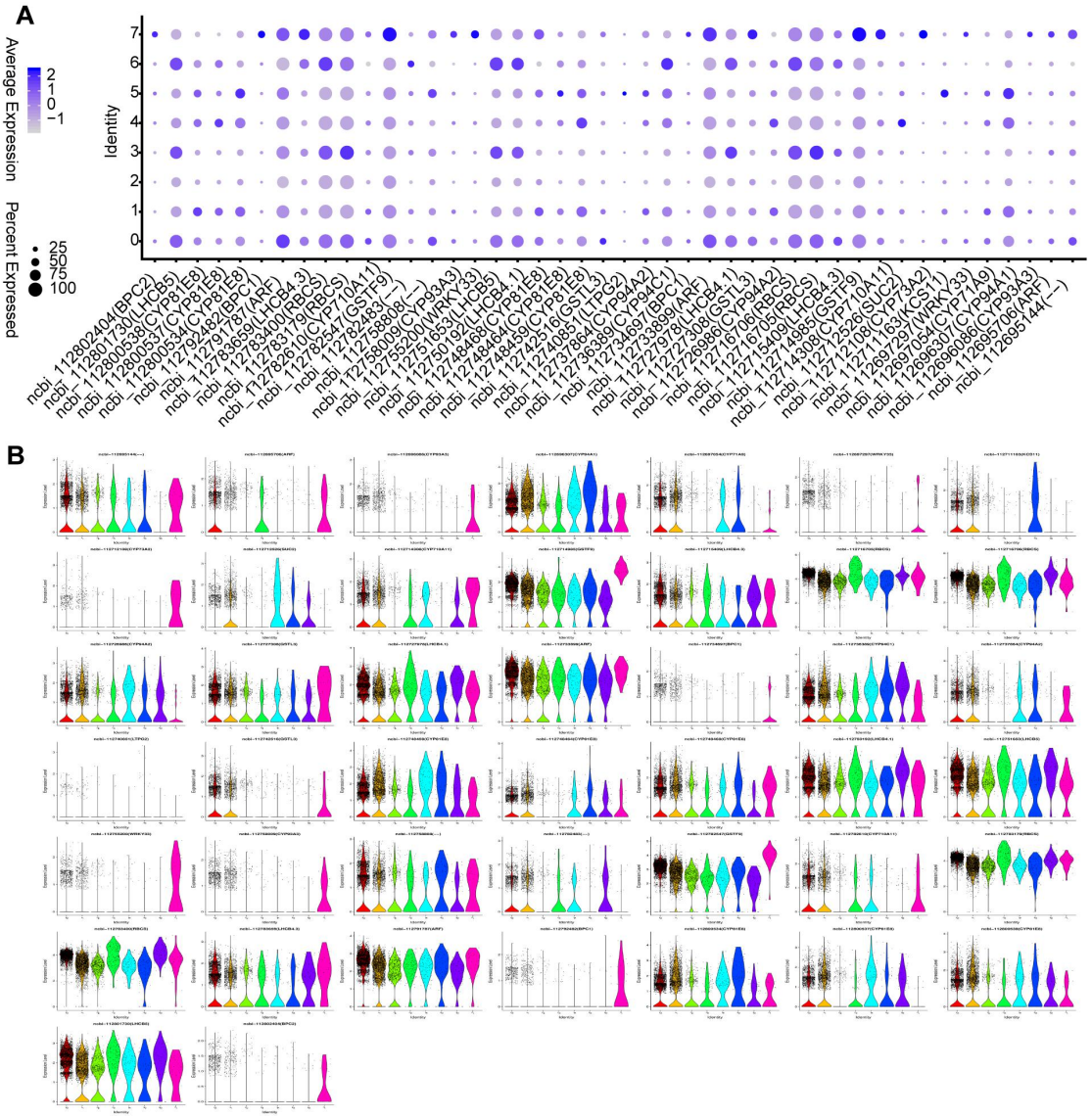

Figure S5. Dot and violin plots of 44 reported marker genes for cell cluster.

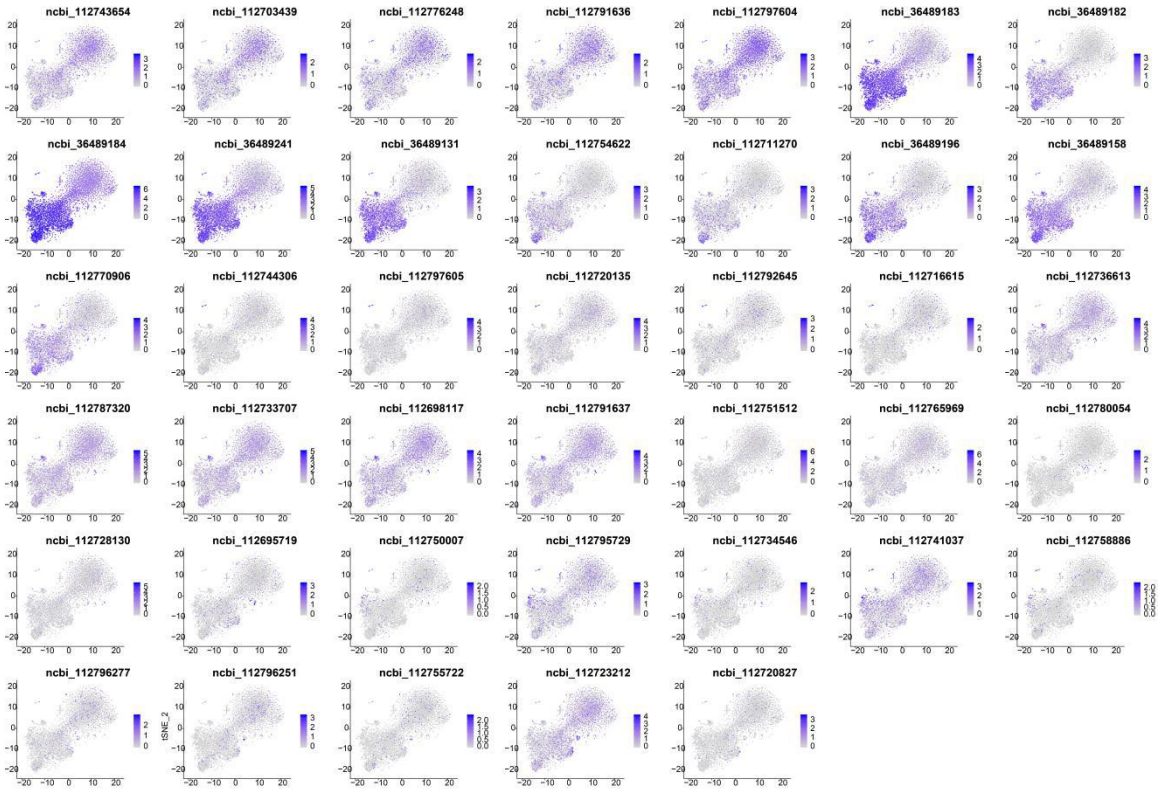

Figure S6. Novel 40 marker gene distribution in t-SNE map, which selected from each cell cluster (Top five genes in each cluster).

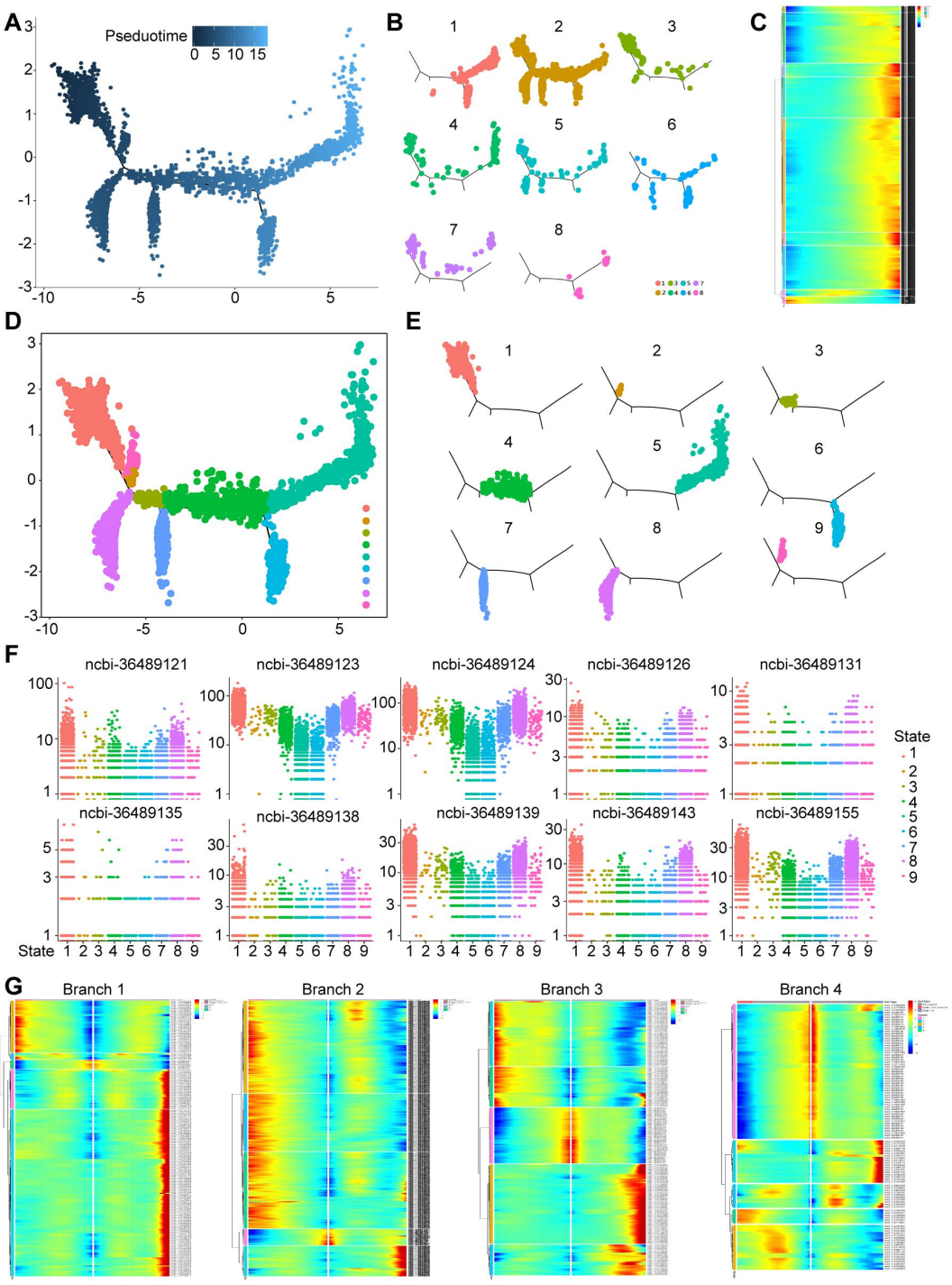

Figure S7. Pseudotime trajectory analysis identified the total trajectory map of all cell. A, development trajectory of all leaf cells. B, the distribution of each separated cell clusters in trajectory map. C, Clustering and expression kinetics of identified DEGs along with the pseudotime trajectory. D-E, All single-cells can be divided into nine state profiles of cell differentiation based on the trajectory analysis. F, Pseudo-time trajectory analysis identified the expression pattern of ten critical marker genes in nine state profiles along with peanut leaf development and differentiation. G, Clustering and expression kinetics of identified DEGs in each cell differentiation branch.

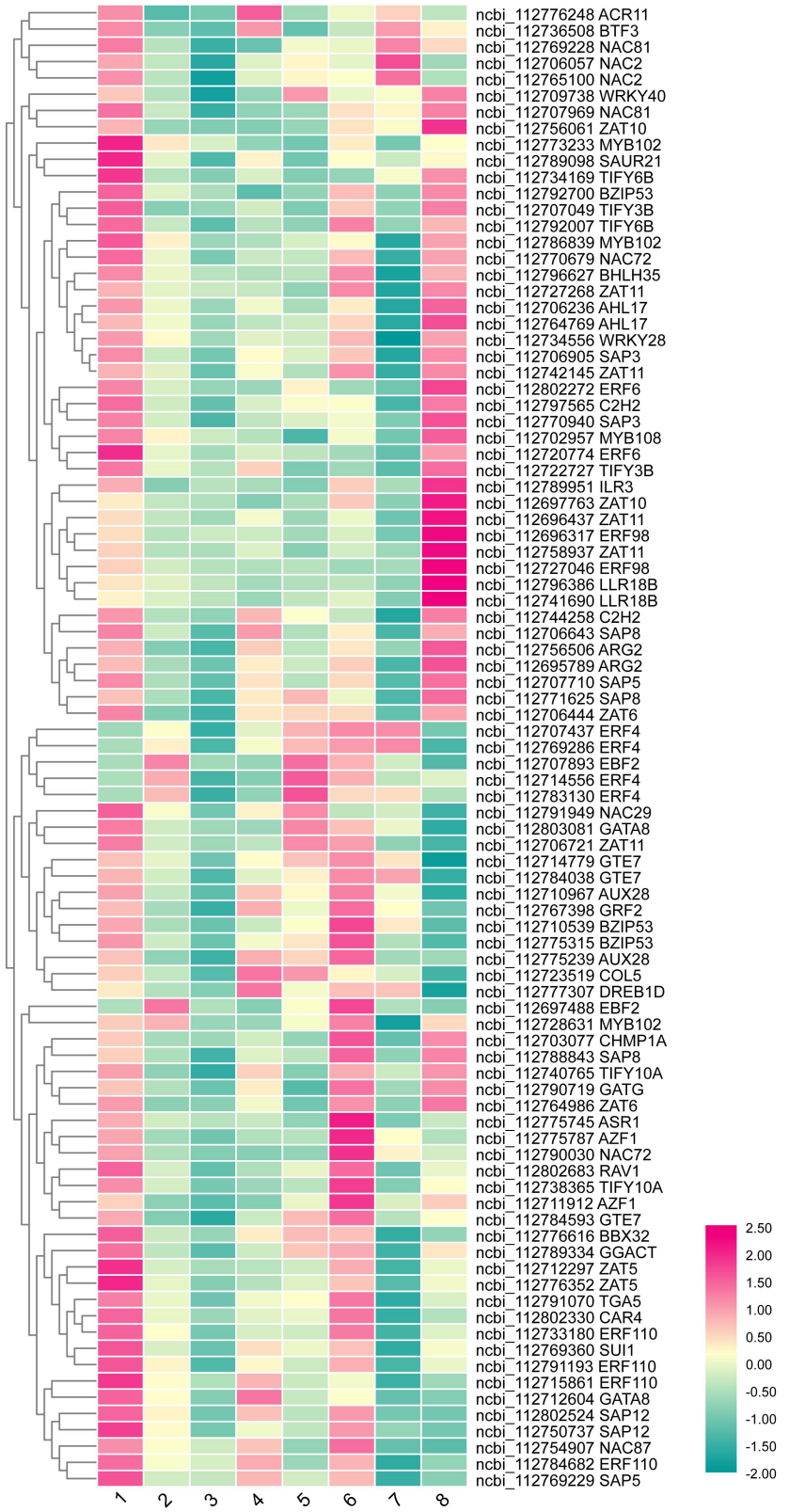

Figure S8. Heatmap displayed the average expression level of represented TFs in branch 2 of all cell trajectory.

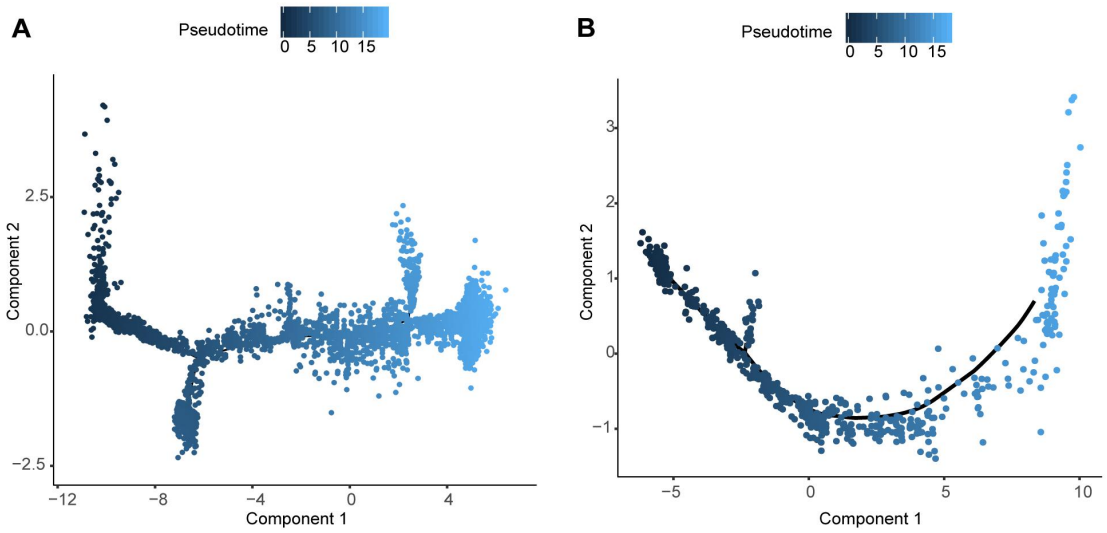

Figure S9. Pseudotime trajectory of mesophyll and epidermal cell development. A, primordium cell developed into mesophyll cell. B, primodium developed into epidermal cell.

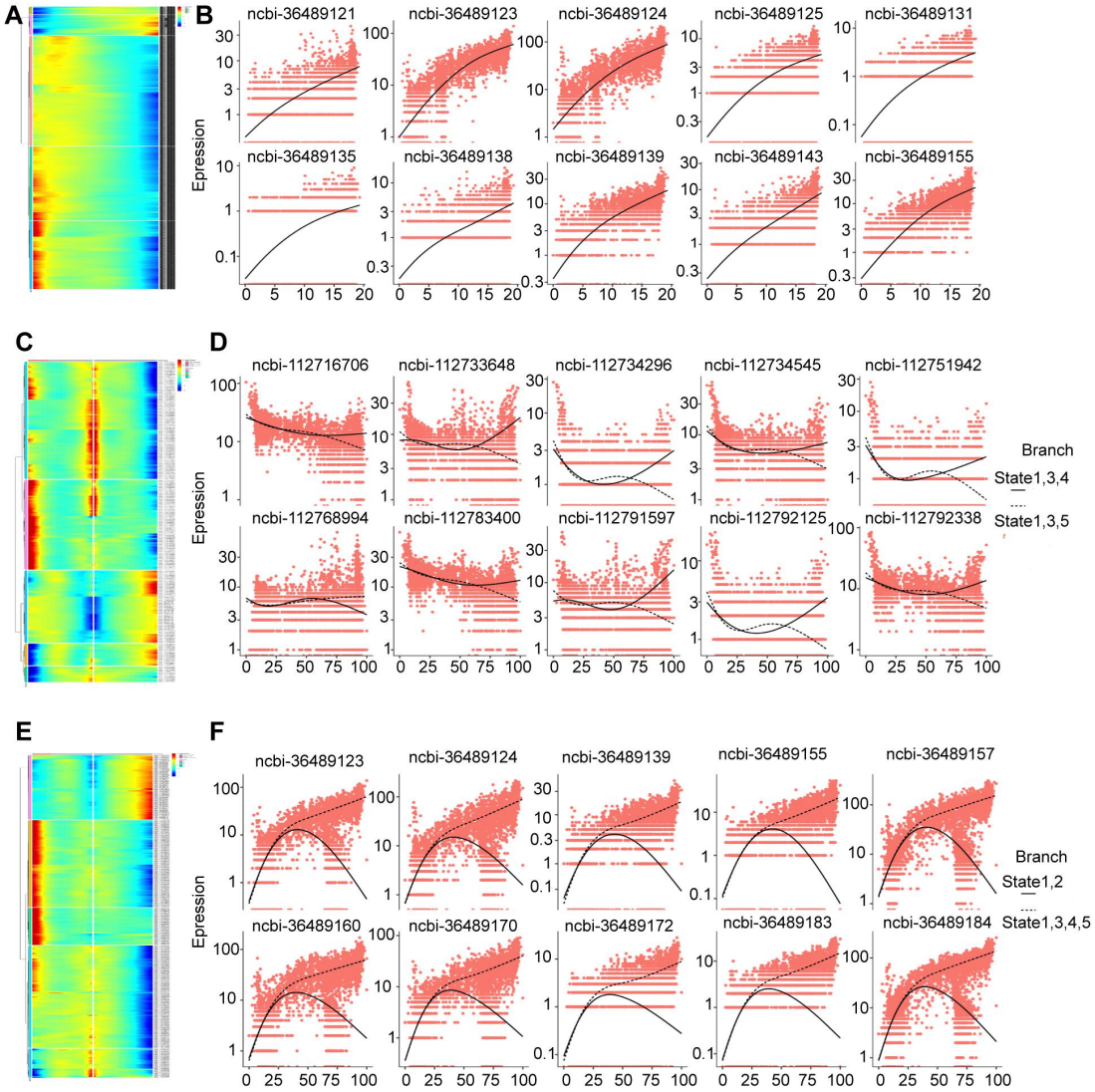

Figure S10. Pseudotime trajectory analysis of primordium and parenchymal developed into mesophyll cell. A, Clustering and expression kinetics of identified DEGs (**Table S13**) along with the main stem of pseudotime trajectory. B, expression distribution of top ten DEGs with highest level. C, Clustering and expression kinetics of identified DEGs in cell differentiation branch 1. D, expression distribution of top ten DEGs at the point 1 of cell differentiation state. E, Clustering and expression kinetics of identified DEGs in cell differentiation branch 2. F, expression distribution of top ten DEGs at the point 2 of cell differentiation state.

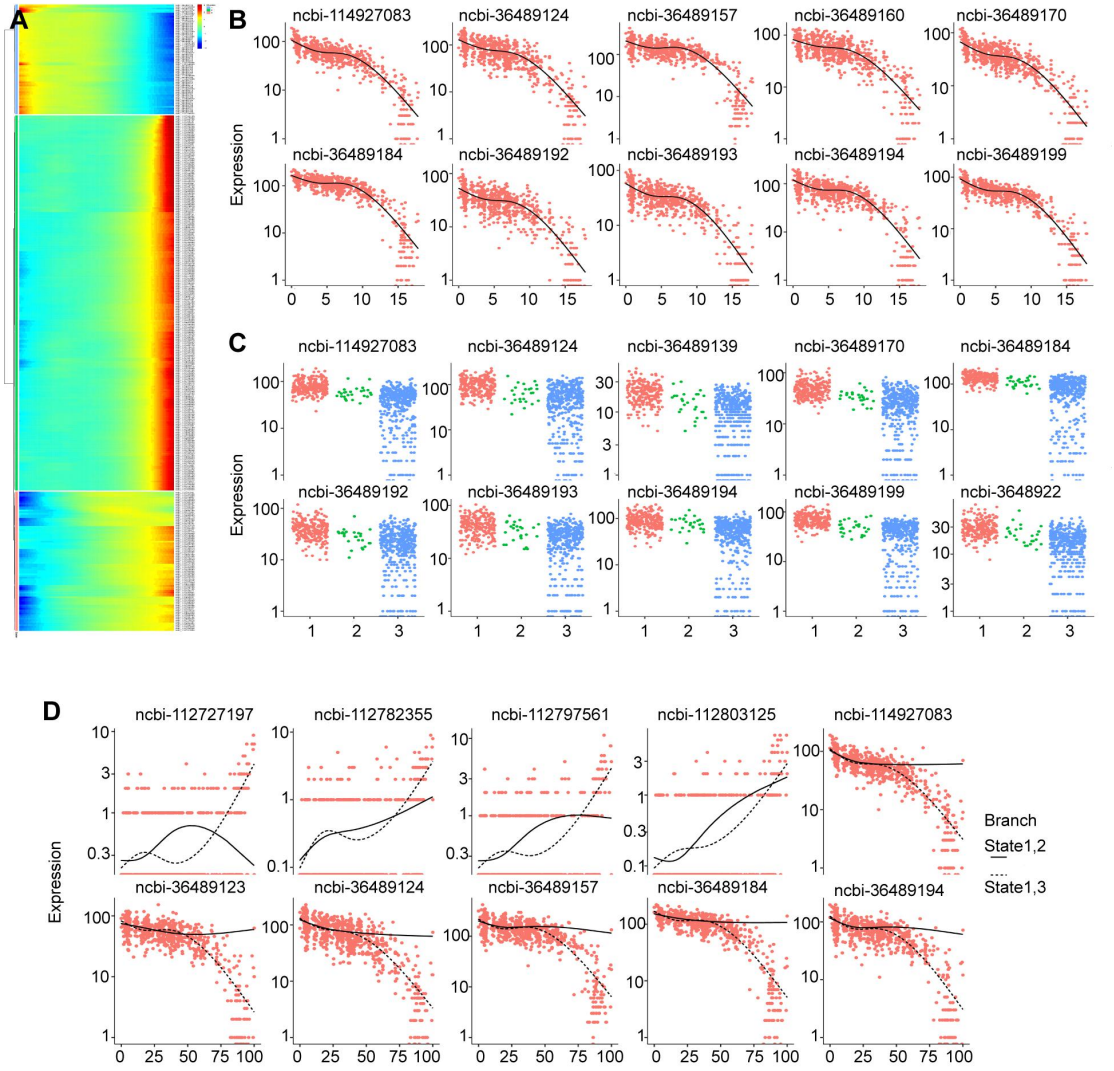

Figure S11. Pseudotime trajectory analysis of primordium developed into epidermal cell. A, Clustering and expression kinetics of identified DEGs (**Table S19**) along with the main stem of pseudotime trajectory. B, expression distribution of top ten DEGs with highest level. C, expression distribution of top ten DEGs correlated to cell differentiation state. D, expression distribution of top ten DEGs at the point 1 of the cell differentiation state.

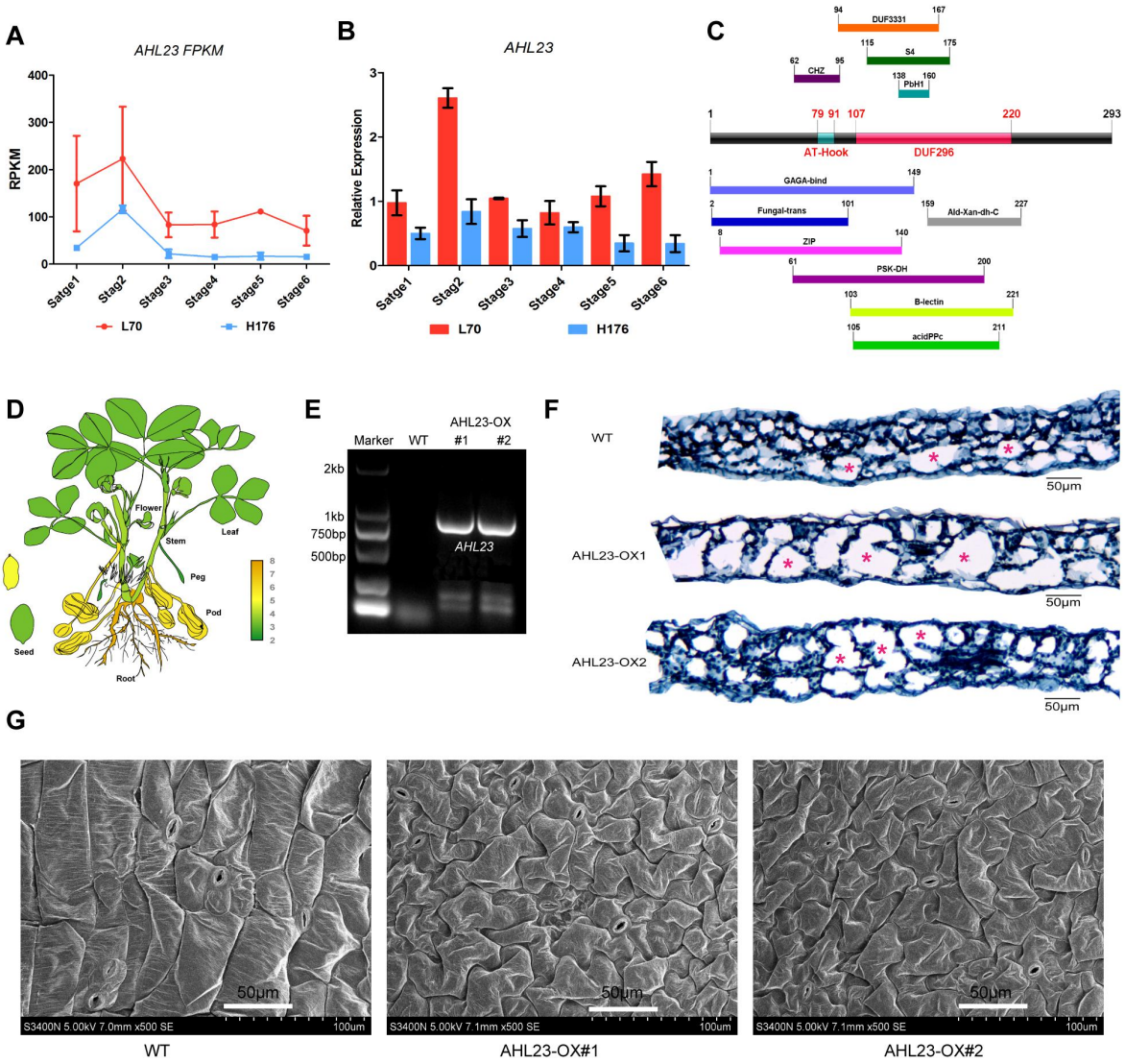

Figure S12. Molecular function analysis of peanut AHL23. A, previous transcriptome (RNA-seq) analysis identified the AHL23 transcript abundance during the seed development of normal oleic acid variety (L70) and high oleic acid variety (H176). B, realt-time PCR validated the truth expression level of AHL23 in L70 and H176 seed. C, SMART database (smart.embl-heidelberg.de) predicted the conserved domain in AHL23 protein. D, eGFP displayed the relative expression of AHL23 in peanut plant, the value of AHL23 transcript abundance downloaded from the Peanutbase.org. E, PCR cloned the AHL23 coding sequence in AHL23-OX lines. F, Transverse sections of leaf blades that collected from wild-type and AHL23-OX seedlings grown at vegetative stage, scale bar 50μm. G, Scanning electron microscope (SEM) observed the size of epidermal cell in WT and AHL23-OX transgenic lines, scale bar 50μm.

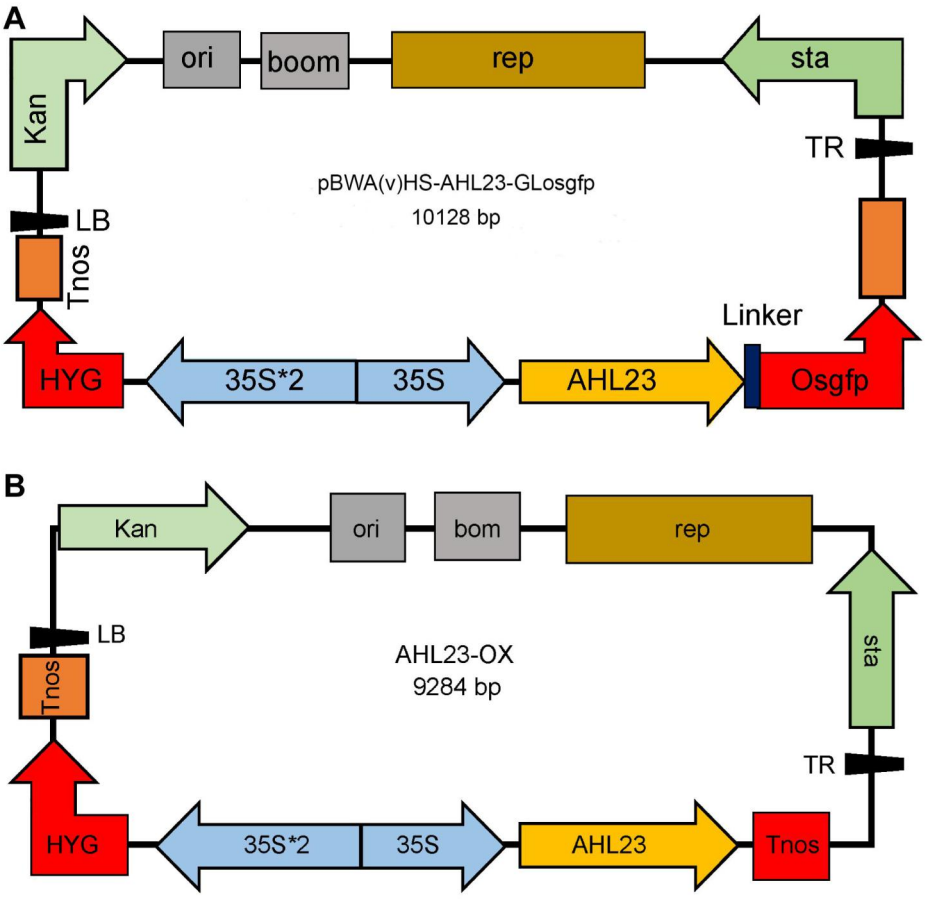

Figure S13. Vector construction for AHL23 function study. A, AHL23-GFP vector for subcellular localization analysis. B, AHL23-OX vector for generation of AHL23-OX transgenic lines.

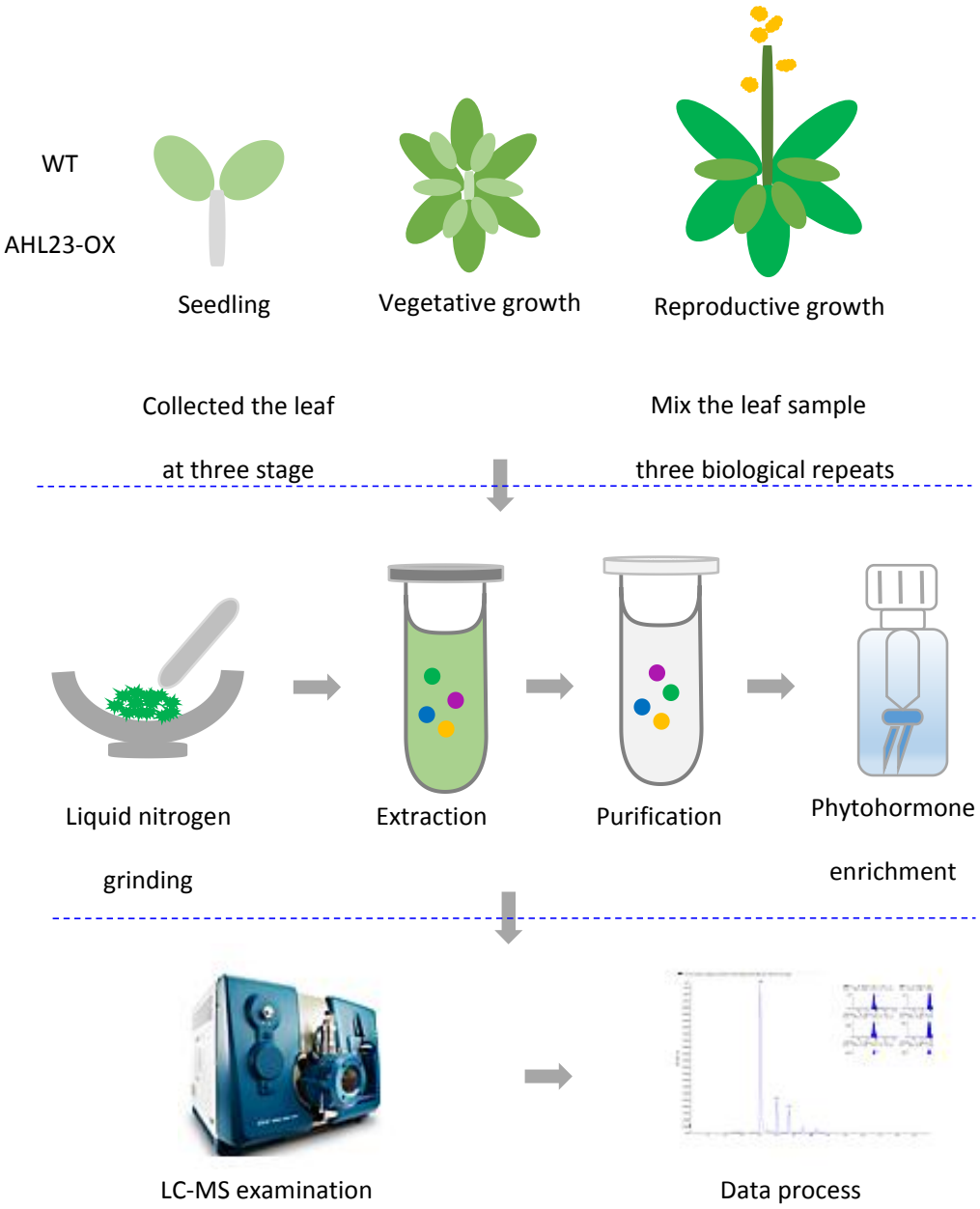

Figure S14. Brief process of phytohormone examination.