


CORRESPONDENCE

Open Access



Response to: “Merit of integrating in situ transcriptomics and anatomical information for cell annotation and lineage construction in single-cell analyses of *Populus*”

Shaoming Liang¹, Yiling Li¹, Yang Chen¹, Heng Huang¹, Sijia Li¹, Yuanzhong Jiang¹ and Tao Ma^{1*} 

This reply refers to the comment available online at <https://doi.org/10.1186/s13059-024-03227-5>.

*Correspondence: matao.yz@gmail.com

¹ Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, Sichuan Zoige Alpine Wetland Ecosystem National Observation and Research Station, College of Life Sciences, Sichuan University, Chengdu, China

Main text

We appreciate the correspondence from Chen et al. [1] regarding four recent studies on single-cell RNA-seq in poplar xylem [2–5]. Over the past 2 years, six research groups have investigated the cell types and differentiation trajectories in poplar woody tissues using single-cell and/or spatial transcriptome techniques [2–7], which provided valuable insights for future research on cambium development and secondary growth in tree species. However, as Chen et al. pointed out, there are discrepancies in the conclusions across these studies, particularly in terms of cell type identification and inferred pseudo-time trajectories. To reach a consensus, it is essential to collect and organize all datasets from multiple sources along with other supporting evidence, such as in situ hybridization, transgenic or reporter lines in poplar, as well as in situ cellular transcriptome and anatomical information as suggested by Chen et al. [1]. In this response, we have gathered raw scRNA-seq data from five studies, excluding the study by Du et al. [7] due to the lack of slices containing spatial information in their data. Since different poplar species and reference genomes were employed in these studies, we integrated their expression matrices based on 1:1 orthologous relationship. This resulted in a final dataset comprising 62,256 cells, with a median of 1730 expressed orthologous genes, which can be divided into 21 cell clusters (Fig. 1A). Subsequently, we conducted a comprehensive analysis of the integrated dataset to try to elucidate the observed discrepancies.

First, we collated markers from these studies [8], and the integrated analysis indicates that plasticity in marker genes expression between samples may partially account for these discrepancies. For example, in two studies [2, 6], two genes related to normal programmed cell death, *XCP1* and *XCP2a/b*, were used as markers to identify vessel cells (Fig. 1B). After data integration, we found that these genes are specifically expressed in cell cluster 13 and showed high consistency across all samples. Similarly, the genes



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

LAC17 and *CESA4*, which were used for identifying xylem cells (including fibers & vessels) in another study [3], were specifically expressed in clusters 0 and 13 in all samples (Fig. 1B). These findings demonstrated the reliability and broad applicability of these markers in the identification of xylem cells. However, according to the AspWood database (plantgenie.org), XCPs may also be expressed in fiber cells, so there is still a lack of more marker genes to distinguish fiber and vessel cells. In comparison, the xylem parenchyma cell marker gene *ABRI*, identified by in situ hybridization in the study of Chen et al. [3], was mainly expressed in clusters 4 and 17 in all samples, except for one study where the gene was generally expressed in most cell types [5] (Fig. 1B). In particular, the marker *WPP2* used for fiber identification in study of Xie et al. [4], and the candidate marker gene Expansin A6 for vessels in study of Tung et al. [5], did not show any cell type-specific expression patterns in almost all samples (Fig. 1B). These results indicate a wide diversity in the expression patterns of marker genes among different studies, which may be affected by factors such as seedling growth status, experimental procedures and differences between species, or more importantly, these genes may not be reliable as molecular markers for cell type identification. Nevertheless, it is undeniable that comparing multiple species and samples will be advantageous in identifying more conserved marker genes for achieving greater accuracy in cell identification in future research.

Second, integrative study confirms a significant level of cell heterogeneity in the cambium region. In line with the findings in the study of Li et al. [6], our results indicated that the previously recognized markers of cambium, *PXY* and *WOX4*, are extensively expressed in diverse cell clusters (Fig. 1B). We also found that the other markers used in these studies to identify cambium cells were not specifically expressed in any particular cell cluster (Fig. 1B). This suggests that the cells composing cambium may have strong cellular heterogeneity, or that these markers are also expressed in adjacent cells of cambium. Therefore, we recommend that these cells be referred to as “cambium region” or “cambium zone” until more definitive markers for cambium cells are established. Furthermore, we found that in the scRNA-seq data of debarked stems, the expression of cambium region marker genes such as *ANT* [9, 10] was observed (Fig. 1B). This suggests the possibility that cambium cells may be left on the wood side during bark peeling. Notably, the thickness of cell walls in tree stem tissue cells varies significantly, which can alter their proportions during protoplast preparation and lead to an inaccurate representation of the actual cellular composition of stem tissue in scRNA-seq data. In comparison, single-nuclei RNA-seq would cover a more comprehensive range of cell types and provide a more precise evaluation of cellular composition.

Third, correlation analysis shows that LCM (laser capture microdissection) combined with RNA-seq can provide another dimension of reference for cell classification, but the resolution is limited. In the study of Tung et al. [5], LCM-RNA-seq data

(See figure on next page.)

Fig. 1 Single-cell atlas integrated based on 1:1 orthologous relationship. **A** UMAP visualization of different samples. Note that the data source is labeled to the left of the UMAP. Chen et al. [3] collected xylem and phloem cells separately, and Li et al. [6] collected primary growth stems and secondary growth stems separately, so we displayed these results separately. **B** The proportion of each cluster and the expression pattern of marker genes in different samples. **C** The correlation between each cluster and LCM-RNA-seq in different samples

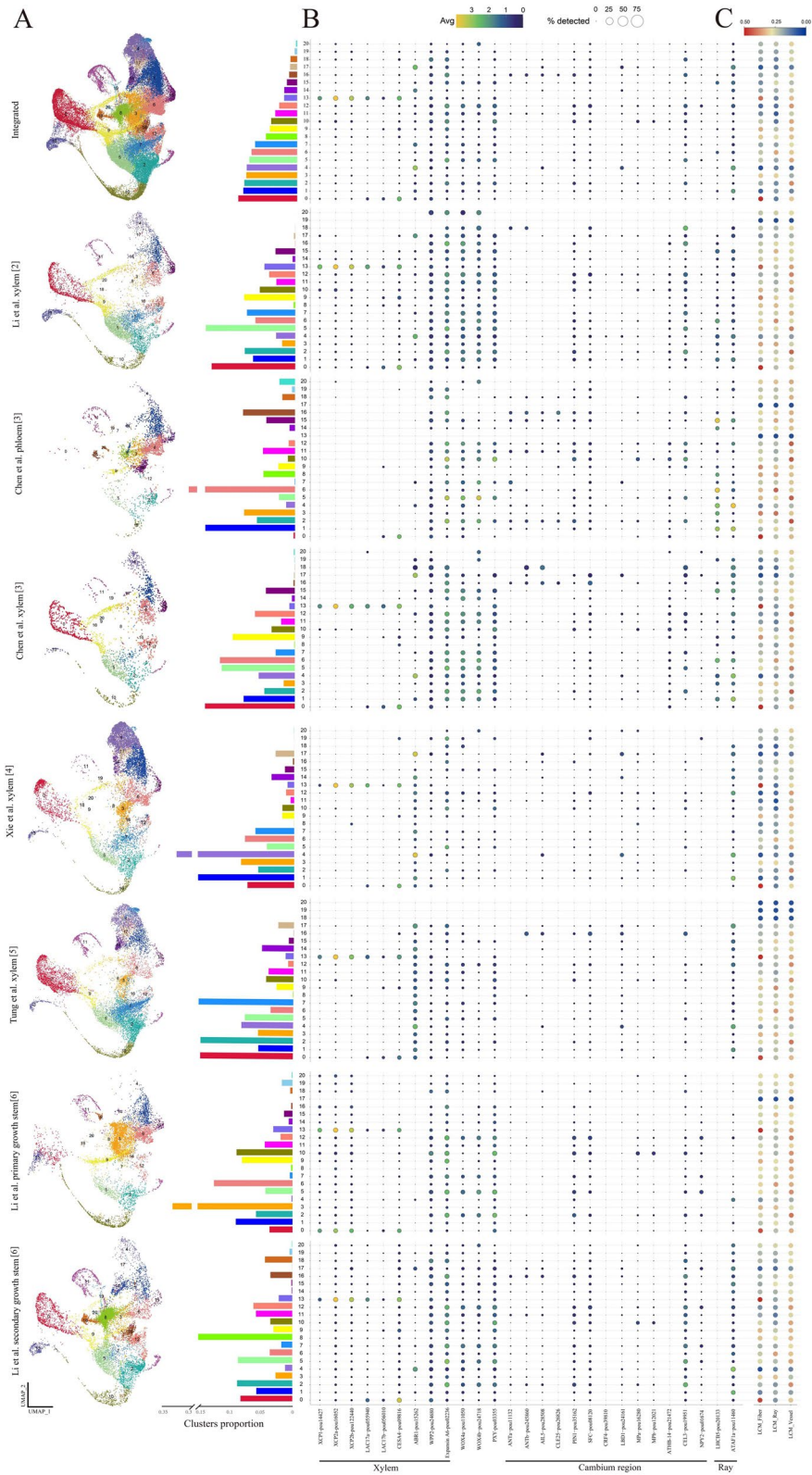


Fig. 1 (See legend on previous page.)

for fiber, vessel, and ray cells were generated. We therefore conducted a correlation analysis between the expression levels of LCM-RNA-seq and the cell clusters identified here. The results showed that LCM-RNA-seq of fibers was closely related to the gene expression in cell clusters 0 and 13, which is generally consistent with the above results of cell annotation by xylem marker genes (Fig. 1C). Furthermore, LCM-RNA-seq of vessels revealed a similar correlation level with various cell types, including clusters 2, 5, 10, and 12 (Fig. 1C). However, there was not enough evidence to support the specific expression of genes related to secondary cell wall biosynthesis in these clusters. These results suggest that LCM-collected vessels may be a mixture of this cell type in multiple transcriptional states and that these clusters may represent vessels at early stages of development. Finally, in cell cluster 6, which had a slight correlation with the LCM-RNA-seq of ray cells (Fig. 1C), we did not find strong expression of the ray markers *LHCBS* and *ATAF1a*, which were validated by RNA in situ hybridization in two studies [2, 6]. This may be due to the significant heterogeneity of gene expression within anatomically defined cell types, and potential biases can be introduced when these cells are artificially combined and analyzed using RNA-seq, which requires further evaluation in the future. Therefore, this result also indicates that whether LCM can accurately distinguish various cell types in xylem remains questionable.

Taken together, our results demonstrate that there are still limitations in using existing techniques for xylem cell classification and subsequent analysis. In particular, identification of cell types based on a few marker genes may cause many conflicts. Therefore, it is still necessary to utilize more xylem single-cell data from different species and developmental stages to mine more reliable and conserved marker genes. In addition, the combination of single-cell, single-nuclei RNA-seq, spatial data, and more refined LCM analysis will also bring us more discoveries. It is believed that with the advancement of technology and the accumulation of relevant evidence, our understanding of woody plants will be significantly improved in the future.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13059-024-03228-4>.

Additional file 1. Review history.

Review history

The review history is available as Additional file 1.

Peer review information

Wenjing She was the primary editor of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Authors' contributions

T.M. conceived the project. S.L. analyzed the data. Y.L., Y.C., H.H., S.L., and Y.J. participated in the figures' design and visualization. T.M. and S.L. wrote the paper. All authors discussed the results and commented on the manuscript. The authors read and approved the final manuscript.

Funding

This work was supported by the National Key Research and Development Program of China (2021YFD2201100), National Natural Science Foundation of China (31922061 and 32271828), and Fundamental Research Funds for the Central Universities (2020SCUNL103).

Availability of data and materials

The data were downloaded from GenBank (PRJNA703312 [2] and PRJNA746591 [5]) and the National Genomics Data Center (PRJCA005543 [3], PRJCA014789 [4] and PRJCA015012 [6]). The single-cell expression matrix integrated is available on Zenodo [11].

Declarations**Ethics approval and consent to participate**

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 14 November 2023 Accepted: 27 March 2024

Published online: 03 April 2024

References

1. Chen YL, Hsieh JW, Kuo SC, Kao CT, Tung CC, Yu JH, Chang TH, Ku C, Xie J, Zhang D, Li Q, Lin YJ. Merit of integrating *in situ* transcriptomics and anatomical information for cell annotation and lineage construction in single-cell analyses of *Populus*. *Genome Biol.* 2024. <https://doi.org/10.1186/s13059-024-03227-5>.
2. Li H, Dai X, Huang X, Xu M, Wang Q, Yan X, Sederoff RR, Li Q, et al. Single-cell RNA sequencing reveals a high-resolution cell atlas of xylem in *Populus*. *J Integr Plant Biol.* 2021;63:1906–21.
3. Chen Y, Tong S, Jiang Y, Ai F, Feng Y, Zhang J, Gong J, Qin J, Zhang Y, Zhu Y, et al. Transcriptional landscape of highly lignified poplar stems at single-cell resolution. *Genome Biol.* 2021;22:319.
4. Xie J, Li M, Zeng J, Li X, Zhang D, et al. Single-cell RNA sequencing profiles of stem-differentiating xylem in poplar. *Plant Biotechnol J.* 2022;20:417–9.
5. Tung CC, Kuo SC, Yang CL, Yu JH, Huang CE, Liou PC, Sun YH, Shuai P, Su JC, Ku C, Lin YJ, et al. Single-cell transcriptomics unveils xylem cell development and evolution. *Genome Biol.* 2023;24:3.
6. Li R, Wang Z, Wang J, Li L, et al. 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.* 2003;4:100665.
7. Du J, Wang Y, Chen W, Xu M, Zhou R, Shou H, Chen J, et al. High-resolution anatomical and spatial transcriptome analyses reveal two types of meristematic cell pools within the secondary vascular tissue of poplar stem. *Mol plant.* 2023;16:209–28.
8. Liang S, Li Y, Chen Y, Huang H, Zhou R, Ma T, et al. Application and prospect of single-cell and spatial omics technologies in woody plants. *Forestry Research.* 2023;3(1).
9. Randall RS, Miyashima S, Blomster T, Zhang J, Elo A, Karlberg A, Immanen J, Nieminen K, Lee J-Y, Kakimoto T, et al. AINTEGUMENTA and the D-type cyclin CYCD3; 1 regulate root secondary growth and respond to cytokinins. *Biology Open.* 2015;4:1229–36.
10. Fu X, Su H, Liu S, Du X, Xu C, Luo K, et al. Cytokinin signaling localized in phloem noncell-autonomously regulates cambial activity during secondary growth of *Populus* stems. *New Phytol.* 2021;230:1476–88.
11. Liang S, et al. 2024. Poplar stem scRNA-seq integrated dataset Zenodo. <https://doi.org/10.5281/zenodo.10691111>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.