PNAS BRIEF REPORT | PLANT BIOLOGY OPEN ACCESS

Cell type–specific cytonuclear coevolution in three allopolyploid plant species

Keren Zhang[a,1], Xueru Zhao[a,1], Yue Zhao[a,1], Zhibin Zhang[a,1], Zhijian Liu[b], Ziyu Liu[a], Yanan Yu[a], Juzuo Li[a], Yiqiao Ma[c], Yuefan Dong[a], Xi Pang[a], Xin Jin[a], Ning Li[a], Bao Liu[a], Jonathan F. Wendel[d], Jixian Zhai[b], Yanping Long[b,2], Tianya Wang[a,2], and Lei Gong[a,2]

Edited by David Baulcombe, University of Cambridge, Cambridge, United Kingdom; received June 28, 2023; accepted August 25, 2023

Cytonuclear disruption may accompany allopolyploid evolution as a consequence of the merger of different nuclear genomes in a cellular environment having only one set of progenitor organellar genomes. One path to reconcile potential cytonuclear mismatch is biased expression for maternal gene duplicates (homoeologs) encoding proteins that target to plastids and/or mitochondria. Assessment of this transcriptional form of cytonuclear coevolution at the level of individual cells or cell types remains unexplored. Using single-cell (sc-) and single-nucleus (sn-) RNAseq data from eight tissues in three allopolyploid species, we characterized cell type–specific variations of cytonuclear coevolutionary homoeologous expression and demonstrated the temporal dynamics of expression patterns across development stages during cotton fiber development. Our results provide unique insights into transcriptional cytonuclear coevolution in plant allopolyploids at the single-cell level.

cytonuclear coevolution | single cell | allopolyploids

Allopolyploidy is an important evolutionary process in plants, involving the combination of two or more divergent parental genomes into a single nucleus (1). Allopolyploid events are commonly observed across angiosperms, including examples such as *Arachis* (peanut), *Gossypium* (cotton), and *Triticum* (wheat). The merger of different nuclear genomes but inheritance of only one set (usually maternal) of progenitor organellar genomes generates potential cytonuclear mismatches. Nuclear genes encoding proteins targeted (2) to mitochondria and plastids (abbreviated as mtNOT and ptNOT, respectively) offer an opportunity to investigate how cytonuclear mismatches are resolved in allopolyploids. Previous studies for various allopolyploid tissues using bulk RNAseq data for whole tissues (most commonly leaves) have characterized transcriptional cytonuclear coevolution, both globally and for ptNOT *rbcS* genes encoding the small subunits of RuBisCo (1, 5-bisphosphate carboxylase/oxygenase). A common but not universally observed result is biased maternal *rbcS* homoeolog expression (3–7). The variability of cytonuclear homoeolog expression patterns is more evident when considering more allopolyploid lineages and whole-genomic mt-/pt-NOT (2).

Notably, the assessment of transcriptional cytonuclear coevolution at the level of individual cells remains unexplored, nor is it clear how this might vary among cell types (8). Advances in single-cell transcriptomic sequencing techniques, at the level of the cell (sc-RNAseq) and nucleus (sn-RNAseq), have made it feasible to characterize and compare the single-cell transcriptome atlas of plant cells, capturing cell type–conserved and cell type–specific expression patterns (8, 9). Here, we present an analysis of cytonuclear coevolution in allopolyploids at the single-cell level, using sc-/sn-RNAseq data from eight tissues in three allopolyploid species. Given the variability in transcriptional cytonuclear coevolution (2), we hypothesized i) that maternal preference is variable among cell types within the same tissues of allopolyploids and ii) that there may be temporal dynamics in this response across different developmental stages of the same cell type. Our findings confirm cell type–specific transcriptional cytonuclear variation in homoeolog expression, which is dynamic in development.

Results and Discussion

We collected or generated sc-/sn-RNAseq data from eight shoot and/or root tissues in peanut (*Arachis hypogaea*), cotton (*Gossypium hirsutum*), and wheat (*Triticum aestivum*) for single-cell cytonuclear coevolutionary analyses (Fig. 1 *A*, *B*, *K*, and *L* data generated in this study; Dataset S1). Based on well-defined cell types (Dataset S1) and partitioning of gene expression values into those either targeting organelles or not targeting organelles (mt-/pt-NOT and non-mt-/pt-NOT; Dataset S2), we characterized and compared the proportions of homoeologs exhibiting biased maternal expression, as well as their extents

Author affiliations: [a]Key Laboratory of Molecular Epigenetics of the Ministry of Education, Northeast Normal University, Changchun, Jilin 130024, China; [b]Department of Biology, School of Life Sciences, Institute of Plant and Food Science, Southern University of Science and Technology, Shenzhen, Guangdong 518055, China; [c]Jilin Academy of Vegetable and Flower Science, Changchun, Jilin 130033, China; and [d]Department of Ecology, Evolution and Organismal Biology, Iowa State University, Ames, IA 50010

Author contributions: B.L., J.F.W., T.W., and L.G. designed research; K.Z., X.Z., Y.Z., Y.M., and T.W. performed research; X.J., J.Z., and Y.L. contributed new reagents/analytic tools; K.Z., Z.Z., Zhijian Liu, Ziyu Liu, Y.Y., J.L., Y.D., X.P., and N.L. analyzed data; B.L. supervision; J.F.W. reviewed and edited the paper; and K.Z., X.Z., and T.W. wrote the paper.

The authors declare no competing interest.

[1]K.Z., X.Z., Y.Z., and Z.Z. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: longyp@sustech.edu.cn, wangty309@nenu.edu.cn, or gongl100@nenu.edu.cn.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2310881120/-/DCSupplemental.

Published September 25, 2023.

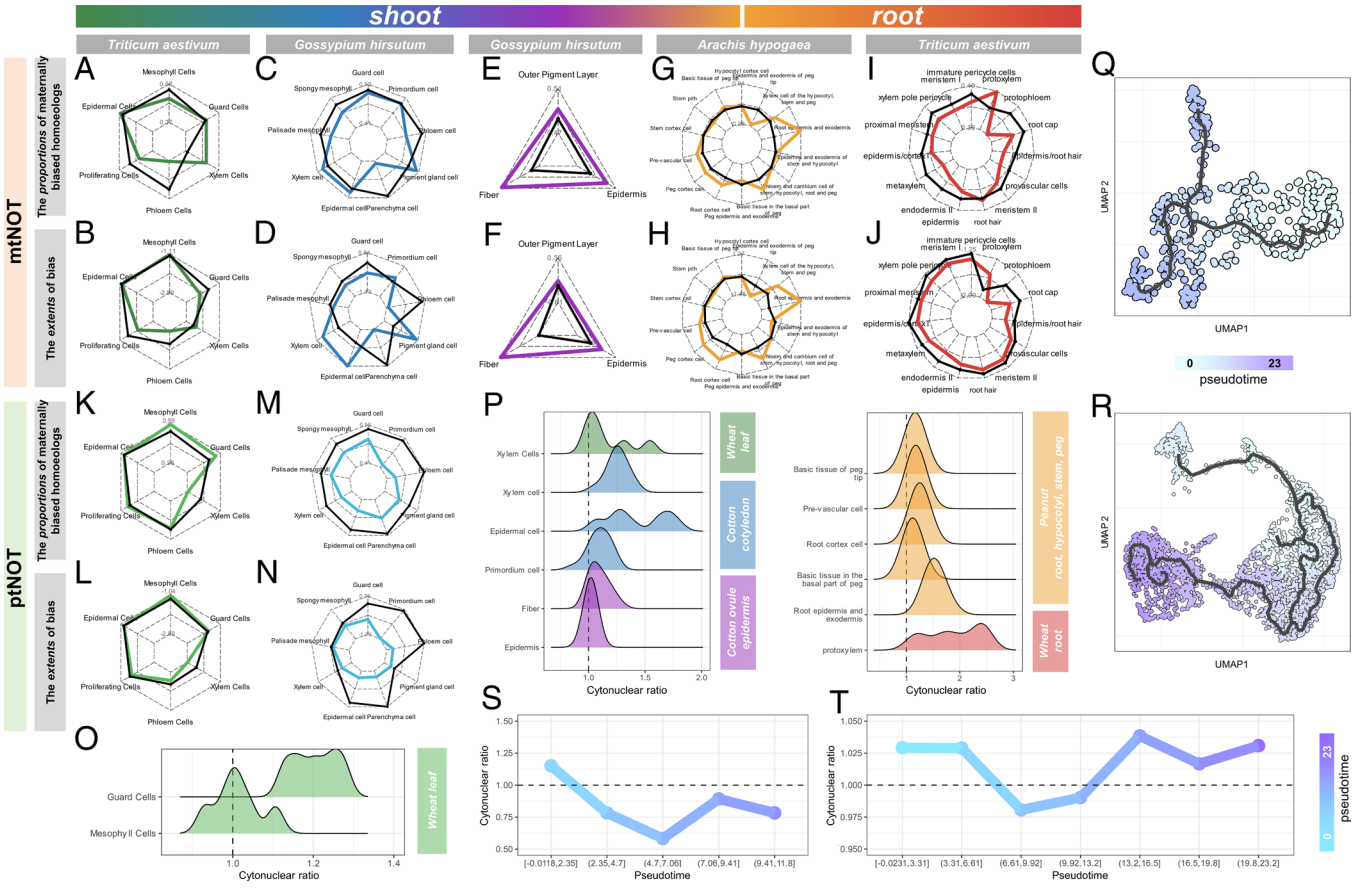

Fig. 1. Maternally biased expression of allopolyploid homoeologs at the single-cell level for cytonuclear genes. (*A*–*J*) Maternally biased homoeologous expression of nuclear genes encoding proteins targeted to mitochondria (mtNOT) (colored lines) vs. other (non-mtNOT) genes (black lines) in different cell types from shoot tissues of wheat (leaf; *A* and *B*), cotton (cotyledon and ovule epidermis; *C*, *D*, *E*, and *F*), and peanut (stem and peg; *G* and *H*), and root tissues of wheat (root; *I* and *J*) and peanut (hypocotyl and root; *G* and *H*). For all cell types specified at the vertices of each radar chart, the proportions of homoeologs exhibiting maternally biased homoeologous expression, and their extents of bias ($\log_2$ transformed ratio of maternal to nonmaternal homoeologous expression) are illustrated in row 1 (*A*, *C*, *E*, *G*, and *I*) and row 2 (*B*, *D*, *F*, *H*, and *J*), respectively. Species with the colored (nonblack) line extend beyond the background (black) line at certain cell-type vertice(s) reflect maternally biased homoeologous expression. (*K*–*N*) Maternally biased homoeologous expression of nuclear genes targeted to the plastids (ptNOT; colored lines) vs. other genes (non-ptNOT; black line) in wheat leaf (*K* and *L*) and cotton cotyledon (*M* and *N*). Illustrated are the proportions of homoeologs showing homoeologous expression bias (*K* and *M*) and the relative extents of bias (*L* and *N*). (*O* and *P*) Density curves of cytonuclear ratios given various thresholds for defining homoeologous expression bias (*O*: ptNOT and *P*: mtNOT; *SI Appendix*). The cell types of different plant tissues are coded using the same color scheme as in *A*–*N* panels. The vertical dashed line denotes equal proportions of homoeologous groups with maternally biased expression in the NOT and non-NOT groups, i.e., 1.0. (*Q* and *R*) Uniform manifold approximation and projection (UMAP) visualization of pseudotemporal development trajectory of cotton fiber as predicted by −1 to 0 dpa (*Q*) and −2 to 2 dpa (*R*) sc-RNA data. The black curve represents the pseudotemporal trajectory, and the dots in gradient colors represent the cotton fiber cells developing at different pseudotime points (cells at earlier stages with smaller pseudotime values are in lighter color). (*S* and *T*) Dynamics of relative homoeologous expression bias as estimated in cytonuclear ratio (*y* axis) in cotton fiber at different pseudotime points (*x* axis) across stages of −1 to 0 dpa (*S*) and −2 to 2 dpa (*T*). The horizontal dashed line denotes equal proportions of homoeologous groups with maternally biased expression in the mtNOT and non-mtNOT groups, i.e., 1.0. Gradient colors are the same as in *Q* and *R*.

of bias (Fig. 1 *A*–*N* and Dataset S4). Additionally, we used the cytonuclear ratio (*Materials and Methods* and *SI Appendix*) to identify cell types harboring an excess of maternally biased homoeologous expression in mt-/pt-NOT groups (ratios higher than 1.0; Fig. 1 *O* and *P*).

Our analyses revealed that not all cell types in the surveyed tissues and species exhibit equivalent cytonuclear coevolutionary responses to allopolyploidy. Specifically, ptNOT homoeologs of mesophyll and guard cells (in wheat leaf; Fig. 1 *K*, *L*, and *O*) exhibited maternally biased expression (i.e., 58.89% and 53.23% of the homoeologs in ptNOT and non-ptNOT groups of mesophyll cells; Fig. 1*K*), whereas mtNOT homoeologs of xylem cells (in wheat leaf; Fig. 1 *A*, *B*, and *P*), xylem, epidermal and primordium cells (in cotton cotyledon; Fig. 1 *C*, *D*, and *P*), and differentiated fiber cells and epidermal cells (in cotton ovule epidermis; Fig. 1 *E*, *F*, and *P*) displayed maternally biased expression. Intriguingly, maternal biased expression was also observed in four peanut tissues, peg tip and basal, cortex cells, root epidermal and exodermal cells, and prevascular cells (Fig. 1 *G*, *H*, and *P*) and protoxylem cells (in wheat root; Fig. 1 *I*, *J*, and *P*). This variation in maternal preference among cell types in several different species highlights the complexities that underlie the transcriptional bias of homoeologs classically generated from bulk RNAseq, which simply aggregates the transcript profile across cell types thus yielding a composite average (8). Although the developmental significance of this variation is unclear, our data point to the need for studies investigating cell type–specific functional relationships between homoeologs with respect to cytonuclear physiology. This line of research is likely to provide certain clues about the potential factors (e.g., organellar function requirement) that generate variation in transcriptional cytonuclear coevolution (10–12).

We also analyzed the temporal dynamics of homoeologous expression bias during cotton fiber development using available sc-RNAseq data from different developmental stages [−1 to 0 days postanthesis (dpa); −2 to 2 dpa]. Based on the −1 to 0 dpa sc-RNA data, we inferred the fiber developmental trajectory and subdivided

the stages of fiber differentiation based on pseudotime values (earlier stages assigned smaller pseudotime values; Fig. 1 *Q* and *S*). Our analysis revealed that mtNOT homoeologs in fiber cells exhibited changing of cytonuclear ratios from higher than 1.0 ((−0.0118,2.35] pseudotime) to lower than 1.0 ((2.35,11.8] pseudotime) as they approach ending anthesis (0 dpa; Fig. 1*S*). This implicates attenuating maternal biased expression in mtNOT groups. We observed a similar decrease of homoeologous expression bias during the early stages of fiber differentiation ((−0.0231,13.2] pseudotime; Fig. 1 *R* and *T*) as revealed by the other −2 to 2 dpa sc-RNA data. These observations coincide with the decrease of nuclear-mitochondrial reactive oxygen species (ROS) scavenging during the fiber differentiation period from −3 to 0 dpa (13). As revealed in later sc-RNA data, a high level of maternal mtNOT homoeologous biased expression (cytonuclear ratios higher than 1.0) when approaching 2 dpa ((13.2,23.2] pseudotime) is also consistent with a previous hypothesis that cotton fiber undergoes a period with high mitochondrial demand before 3 dpa (14). These data may thus hint at a potential causal relationship between dynamic maternal mtNOT homoeologous biased expression and mitochondrial demand during fiber development.

Our study represents an investigation of cytonuclear coevolution in allopolyploids at the single-cell level. Given different homoeologous expression patterns among cell types, yet constant DNA composition, our data implicate developmentally sensitive transcriptional machinery and epigenomic mechanisms as being responsible for this form of homoeologous expression variation in the same species. Our findings also indicate that transcriptional cytonuclear coevolution may more precisely be studied at the single-cell level, rather than as an aggregated metric across multiple cell types. Future single-cell homoeologous expression analyses using specific tissues at certain developmental stages in allopolyploids having different parental divergences and/or ages of polyploidy events may provide insights into any cytonuclear coevolutionary generalities and/or uniqueness at the single-cell level.

Materials and Methods

Public sc-/sn-RNAseq data were used from peanut (stem, peg, hypocotyl, and root), cotton (cotyledon and ovule epidermis), and common wheat (root; cultivar Aikang 58), in addition to 10× genomics sn-RNAseq data generated from 7-d-old wheat seedling leaves (cultivar CS; Dataset S1) for this study. For the public data, we used their published cell-type definitions and expression matrices. Seurat was used for sample clustering and cell identity annotation for the wheat seedling leaves (Dataset S3). OrthoFinder and pSONIC were employed to partition expression of duplicate genes into their respective homoeologs. Transcriptional cytonuclear coevolution was evaluated by quantitatively analyzing the proportions of homoeologous groups that showed maternal-biased expression in mt-/pt-NOT vs. that of non-NOT and qualitatively estimating the extent of bias for both mt-/pt-NOT and non-NOT homoeologs. The "cytonuclear ratio" was calculated as the proportion of maternal-biased expression homoeologous groups in the NOT group compared to those in the non-NOT group. Cell types exhibiting cytonuclear ratios >1.0 across bias extent categories (various thresholds for defining homoeologous expression bias) were identified. Monocle3 was employed to reconstruct pseudotemporal trajectories for two cotton fiber libraries at different developmental stages. Detailed information about the materials and methods adopted is described in *SI Appendix*.

Data, Materials, and Software Availability. Sequencing data have been deposited in NCBI (PRJNA985908) (15). All other data are included in the manuscript and/or supporting information. Previously published data were used for this work (16–20).

ACKNOWLEDGMENTS. This work was supported by "the Fundamental Research Funds for the Central Universities (2412023YQ005)" and "Chinese Academy of Sciences (CAS) Youth Interdisciplinary Team (JCTD-2022-06)". We would like to thank Wei Gao and Guowei Li for sharing the cotton cotyledon and peanut sc-RNAseq data used in this study.

1. P. S. Soltis, D. E. Soltis, The role of hybridization in plant speciation. *Annu. Rev. Plant Biol.* **60**, 561–588 (2009).
2. C. E. Grover *et al.*, Variation in cytonuclear expression accommodation among allopolyploid plants. *Genetics* **222**, iyac118 (2022).
3. S. Shan *et al.*, Transcriptome dynamics of the inflorescence in reciprocally formed allopolyploid Tragopogon miscellus (Asteraceae). *Front. Genet.* **11**, 888 (2020).
4. L. Gong, M. Olson, J. F. Wendel, Cytonuclear evolution of rubisco in four allopolyploid lineages. *Mol. Biol. Evol.* **31**, 2624–2636 (2014).
5. D. B. Sloan, Using plants to elucidate the mechanisms of cytonuclear co-evolution. *New Phytol.* **205**, 1040–1046 (2015).
6. R. S. Burton, R. J. Pereira, F. S. Barreto, Cytonuclear genomic interactions and hybrid breakdown. *Annu. Rev. Ecol. Evol. Syst.* **44**, 281–302 (2013).
7. T. Sehrish, V. V. Symonds, D. E. Soltis, P. S. Soltis, J. A. Tate, Cytonuclear coordination is not immediate upon allopolyploid formation in Tragopogon miscellus (Asteraceae) allopolyploids. *PLoS One* **10**, e0144339 (2015).
8. K. H. Ryu, Y. Zhu, J. Schiefelbein, Plant cell identity in the era of single-cell transcriptomics. *Annu. Rev. Genet.* **55**, 479–496 (2021).
9. K. Xia *et al.*, The single-cell stereo-seq reveals region-specific cell subtypes and transcriptome profiling in Arabidopsis leaves. *Dev. Cell* **57**, 1299–1310.e4 (2022).
10. E. Winter-Sluiter, A. Läuchli, D. Kramer, Cytochemical localization of K+-stimulated adenosine triphosphatase activity in Xylem parenchyma cells of barley roots. *Plant Physiol.* **60**, 923–927 (1977).
11. F. B. P. Wooding, Absorptive cells in protoxylem: Association between mitochondria and the plasmalemma. *Planta* **84**, 235–238 (1969).
12. T. Ren, S. M. Weraduwage, T. D. Sharkey, Prospects for enhancing leaf photosynthetic capacity by manipulating mesophyll cell morphology. *J. Exp. Bot.* **70**, 1153–1165 (2019).
13. K. Liu *et al.*, Comparative proteomic analysis reveals the mechanisms governing cotton fiber differentiation and initiation. *J. Proteomics* **75**, 845–856 (2012).
14. G. N. Thyssen, X. Song, M. Naoumkina, H.-J. Kim, D. D. Fang, Independent replication of mitochondrial genes supports the transcriptional program in developing fiber cells of cotton (Gossypium hirsutum L.). *Gene* **544**, 41–48 (2014).
15. K. Zhang, Cell type-specific transcriptional cytonuclear coordination in allopolyploid plants. *NCBI Bioproject*. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA985908. Deposited 21 June 2023.
16. L. Zhang *et al.*, Asymmetric gene expression and cell-type-specific regulatory networks in the root of bread wheat revealed by single-cell multiomics analysis. *Genome Biol.* **24**, 65 (2023).
17. Y. Qin *et al.*, Single-cell RNA-seq reveals fate determination control of an individual fibre cell initiation in cotton (Gossypium hirsutum). *Plant Biotechnol. J.* **20**, 2372–2388 (2022).
18. Y. Liu *et al.*, Spatial transcriptome analysis on peanut tissues shed light on cell heterogeneity of the peg. *Plant Biotechnol. J.* **20**, 1648–1650 (2022).
19. L. Long *et al.*, Single-cell transcriptome atlas identified novel regulators for pigment gland morphogenesis in cotton. *Plant Biotechnol. J.* **21**, 1100–1102 (2023).
20. D. Wang *et al.*, Cell-specific clock-controlled gene expression program regulates rhythmic fiber cell growth in cotton. *Genome Biol.* **24**, 49 (2023).