

From sugarcane to polyploid crops: Graph pangenomes usher in a new era of complex-genome breeding

Polyploidy is a major driver of plant evolution and a defining feature of many domesticated crops. Comparative cytogenetic and phylogenomic surveys indicate that roughly one-third of extant vascular plant species are polyploids, and recurrent polyploidization has contributed to the domestication and improvement of wheat, cotton, potato, sugarcane, and numerous other staples. However, the multiple, often heterogeneous chromosome complements in polyploids make their genomes substantially more difficult to analyze than those of diploids, hindering efforts to resolve homoeologous chromosomes, quantify gene dosage, and map agronomic traits. Sugarcane represents an extreme case: modern *Saccharum* hybrids originated from interspecific crosses between *Saccharum officinarum* and *Saccharum spontaneum*, followed by extensive and uneven polyploidization and aneuploidy, yielding one of the most structurally complex genomes among major crops (Healey et al., 2024). Conventional linear reference genomes, which perform well in diploids, collapse haplotype diversity, obscure structural and dosage variation, and are poorly suited to such mixed-ploidy architectures. This limitation has contributed to the lag in both basic genomics and breeding applications in highly polyploid crops.

Early efforts in sugarcane genomics date back to the 1990s, when Brazilian research groups initiated expressed-sequence-tag-based studies. Nevertheless, owing to the extraordinary complexity of the sugarcane genome, coupled with the limitations of sequencing technologies at the time, progress remained limited for nearly two decades. A major milestone was achieved in 2018 with the generation of the first allele-resolved genome of the autopolyploid wild species *S. spontaneum* AP85-441 (Figure 1), which provided a foundational framework for genomic analyses in sugarcane and other complex polyploid crops (Zhang et al., 2018). This was followed by gap-free assemblies of related genera (Wang et al., 2023) and studies on the evolution of the *Saccharum* complex (Wang et al., 2025, 2026), which together established a robust evolutionary framework, clarifying chromosome base-number shifts, recent chromosome reduction, and lineage divergence, and providing a scaffold for systematic gene mining and polyploid inheritance analyses in this complex genome (Zhang et al., 2022). Building on these resources, integrative approaches combining population-scale resequencing, transcriptomics, and functional genomics have been applied to dissect key biological traits in polyploid backgrounds, including photosynthetic divergence, sugar accumulation, and stress responses (Zhang et al., 2021; Jiang et al., 2023). For modern cultivated sugarcane, the integration of PacBio HiFi sequencing, Bionano optical mapping, chromatin interaction data, and genetic linkage maps enabled the construction of a

~5.04 Gb partial chromosome-scale assembly of the modern cultivar R570, which is considered a model breeding line because of its resistance to brown rust (*Puccinia melanocephala*). This assembly revealed extensive genome collapse attributable to high homology among chromosome sets derived from *S. officinarum* progenitors and facilitated population genomic analyses of domestication footprints across *Saccharum* wild germplasm (Healey et al., 2024; Garsmeur et al., 2025). Parallel efforts also produced a chromosome-level assembly of the modern cultivar ZZ1, elucidating its complex pedigree composition (Bao et al., 2024). Notably, a high-quality chromosome-scale assembly of the highly allo-autopolyploid modern cultivar XTT22 extended these advances to elite breeding germplasm (Zhang et al., 2025). This work enabled precise analyses of allele dosage, subgenome interactions, and genotype–phenotype relationships, but also exposed the limitations of single-reference frameworks for capturing dosage-sensitive variation.

Building on this trajectory, a recent study led by Zhang and co-workers takes a decisive conceptual step by introducing a multi-scale, polyploid-aware graph pangenome for sugarcane (Huang et al., 2026), aligning with the broader emergence of pangenome graphs in biodiversity genomics. This work is highlighted in a *Science* perspective by Pamela S. Soltis and Douglas E. Soltis as representing the state of the art in plant pangenomics, underscoring its broader relevance beyond a single crop system. By integrating 47–57 haplotypes from nine high-quality assemblies spanning four *Saccharum* species into a unified graph coordinate system, the authors move beyond any single linear reference. The resulting graph captures more than twice the genetic diversity represented by an individual genome, enabling reference-unbiased alignment and markedly improved representation of allelic and structural variation across mixed-ploidy levels. Crucially, dosage information is natively encoded, enabling genome-wide analyses that explicitly consider copy-number differences among homoeologous and paralogous loci.

A distinctive strength of this work lies in its multiscale design, which interconnects genome-, gene-, and protein-level graphs. This architecture allows tracking of haplotype evolution within homoeologous gene systems and supports polyploid-specific analyses that were previously impractical. The authors implement dosage-sensitive genome-wide association studies, uncovering convergent selection on carbohydrate metabolism and cell wall pathways and pinpointing dosage-dependent loci underlying key agronomic traits. Functional validation of *SaTB1* demonstrates how graph-based discoveries can be translated into

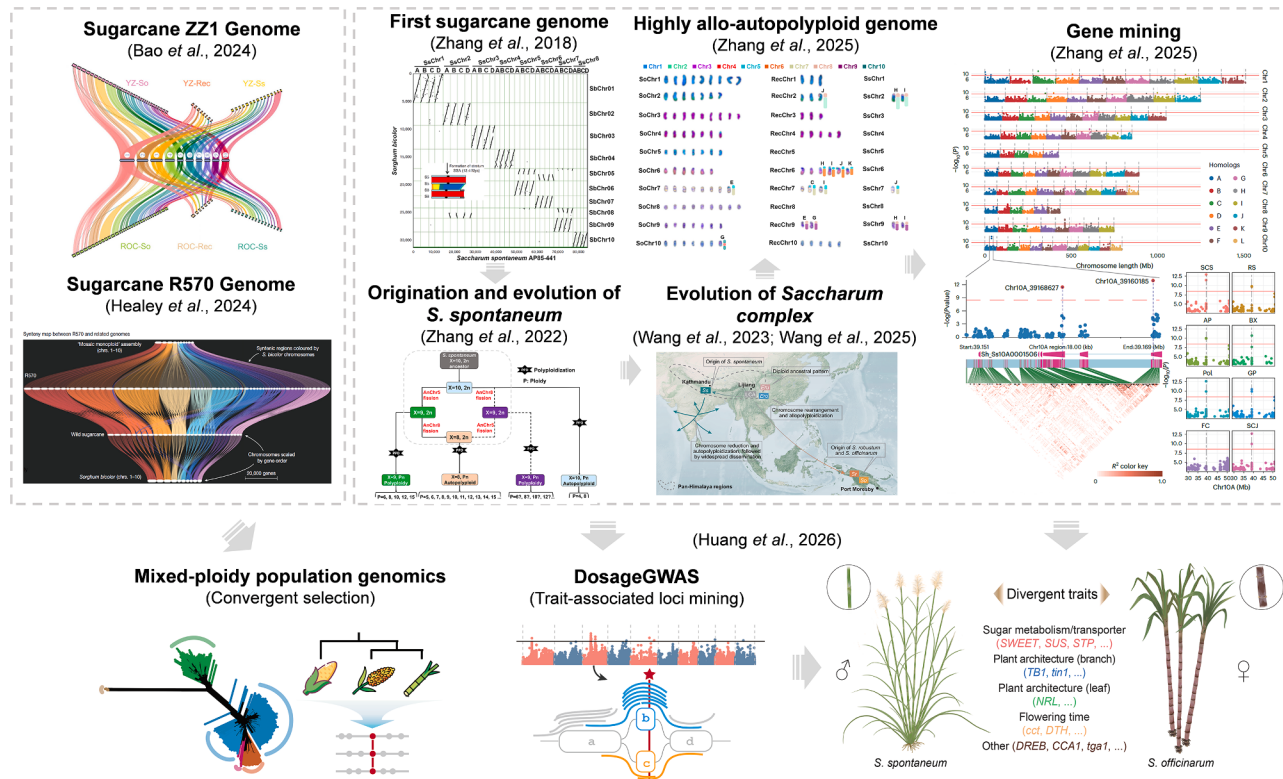


Figure 1. Hierarchical pan-genome integration and transferable polyploid strategies illuminate sugarcane genome diversification.

This figure systematically delineates the research trajectory underlying “multiscale pangenome graphs empower the genomic dissection of mixed-ploidy sugarcane species.” The framework originates from the first reference genome of sugarcane, which established an initial genomic foundation for this exceptionally complex crop (Zhang et al., 2018). Subsequent work resolved the evolutionary history of the wild progenitor *S. spontaneum*, elucidating patterns of polyploid genome evolution, chromosome restructuring, and homoeolog differentiation (Zhang et al., 2022). These efforts were further extended to polyploid relatives within the *Saccharum* complex, revealing both shared and lineage-specific genomic features and refining the evolutionary relationships among closely related species (Wang et al., 2023, 2025). Building on this evolutionary and comparative framework, high-quality linear genome assemblies of modern cultivated sugarcane were generated, enabling precise gene annotation and in-depth mining of functional loci underlying key agronomic traits (Zhang et al., 2025). By integrating these resources with published genomes of modern cultivars, including R570 (Healey et al., 2024) and ZZ1 (Bao et al., 2024), as well as representative assemblies spanning the full *Saccharum* lineage, a multiscale, multidimensional pangenome graph was constructed. This graph-based framework captures genome-, gene-, and protein-level variation across mixed-ploidy levels and establishes a transferable methodological paradigm for polyploid genome analysis (Huang et al., 2026). Leveraging this integrated genomic infrastructure, downstream analyses encompassing epigenomics, proteomics, and trait mapping were conducted, alongside the development of specialized analytical approaches. Together, these multilevel validations provide a comprehensive view of the structural, allelic, and dosage-driven diversification inherent in the polyploid sugarcane genome.

yield-relevant phenotypes, advancing sugarcane genomics from descriptive cataloging of complexity toward predictive, design-oriented breeding.

Beyond sugarcane, this study establishes a broadly generalizable paradigm for research on complex polyploid crops. Its core framework—haplotype-resolved genome assemblies, graph-based integration of structural and dosage variation, and functional validation in a breeding context—is readily transferable to other allopolyploid and autopolyploid species, such as wheat, alfalfa, and potato. By addressing shared core hurdles in polyploid genomics, particularly dosage effects and homoeolog resolution, this work recasts sugarcane from an outlier “ultimate challenge” into a model system for next-generation polyploid crop improvement.

More broadly, this body of work exemplifies how sustained, strategic investment in a single, challenging crop can yield

methodological advances of global relevance. At a time when sequencing technologies, graph-based genome representations, and the urgent need for climate-resilient, high-yield crops are converging, sugarcane genomics has shifted from following the field to helping define it for polyploid species. The sugarcane graph pangenome framework not only accelerates molecular breeding in this economically important crop but also helps unlock the broader potential of polyploidy to enhance yield, resilience, and resource-use efficiency across agriculture, underscoring the pivotal role of sugarcane as a polyploid model for advancing fundamental plant genomics and crop design.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (32300503 to C.W.) and the National Key Research and Development Program of the Ministry of Science and Technology of China (2022YFF1003301 to B.H.).

ACKNOWLEDGMENTS

The authors declare no conflicts of interest.

Received: February 9, 2026

Revised: April 3, 2026

Accepted: April 10, 2026

Published: April 14, 2026

Changsheng Wang and Bin Han*

National Center for Gene Research, State Key Laboratory of Plant Trait Design, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China

*Correspondence: Bin Han (bhan@ncgr.ac.cn)

<https://doi.org/10.1016/j.xplc.2026.101852>

REFERENCES

- Bao, Y., Zhang, Q., Huang, J., Zhang, S., Yao, W., Yu, Z., Deng, Z., Yu, J., Kong, W., Yu, X., et al. (2024). A chromosomal-scale genome assembly of modern cultivated hybrid sugarcane provides insights into origination and evolution. *Nat. Commun.* **15**:3041. <https://doi.org/10.1038/s41467-024-47390-6>.
- Garsmeur, O., Rio, S., Pompidor, N., Lipzen, A., Hervouet, C., Durand, T., Daum, C., Yoshinaga, Y., Butterfield, M., Sanchez, A., et al. (2025). The genomic footprints of wild *Saccharum* species trace domestication, diversification, and modern breeding of sugarcane. *Cell* **188**:7252–7266.e15.
- Healey, A.L., Garsmeur, O., Lovell, J.T., Shengquiang, S., Sreedasyam, A., Jenkins, J., Plott, C.B., Piperidis, N., Pompidor, N., Llaca, V., et al. (2024). The complex polyploid genome architecture of sugarcane. *Nature* **628**:804–810. <https://doi.org/10.1038/s41586-024-07231-4>.
- Huang, Y., Zhang, Y., Zhang, Q., Zhuang, G., Li, C., Wang, B., Gao, R., Xu, Y., Qi, Y., Hua, X., et al. (2026). Multiscale pangenome graphs empower the genomic dissection of mixed-ploidy sugarcane species. *Science* **391**:eadx1616. <https://doi.org/10.1126/science.adx1616>.
- Jiang, Q., Hua, X., Shi, H., Liu, J., Yuan, Y., Li, Z., Li, S., Zhou, M., Yin, C., Dou, M., et al. (2023). Transcriptome dynamics provides insights into divergences of the photosynthesis pathway between *Saccharum officinarum* and *Saccharum spontaneum*. *Plant J.* **113**:1278–1294. <https://doi.org/10.1111/tpj.16110>.
- Wang, B., Zhang, Z., Qi, Y., Wei, Y., Mei, J., Liu, H., Wang, Y., Fang, Y., Hua, X., Ding, H., et al. (2025). Near-complete genome assembly of allotetraploid *Erianthus rockii* reveals unique chromosome evolution and lineage-divergence trajectories in the *Saccharum* complex. *Plant Commun.* **6**:101464. <https://doi.org/10.1016/j.xplc.2025.101464>.
- Wang, T., Wang, B., Hua, X., Tang, H., Zhang, Z., Gao, R., Qi, Y., Zhang, Q., Wang, G., Yu, Z., et al. (2023). A complete gap-free diploid genome in *Saccharum* complex and the genomic footprints of evolution in the highly polyploid *Saccharum* genus. *Nat. Plants* **9**:554–571. <https://doi.org/10.1038/s41477-023-01378-0>.
- Wang, Y., Qi, Y., Zhang, Y., Wang, B., Sun, X., Ding, H., Xu, J., Zhang, Q., and Zhang, J. (2026). Evolution of paleo-duplicated chromosome pairs with enrichment of NBS-LRR genes generated through ρ -WGD in Poaceae. *New Phytol.* **249**:3104–3119. <https://doi.org/10.1111/nph.70862>.
- Zhang, J., Zhang, X., Tang, H., Zhang, Q., Hua, X., Ma, X., Zhu, F., Jones, T., Zhu, X., Bowers, J., et al. (2018). Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum* L. *Nat. Genet.* **50**:1565–1573. <https://doi.org/10.1038/s41588-018-0237-2>.
- Zhang, J., Qi, Y., Hua, X., Wang, Y., Wang, B., Qi, Y., Huang, Y., Yu, Z., Gao, R., Zhang, Y., et al. (2025). The highly allo-autopolyploid modern sugarcane genome and very recent allopolyploidization in *Saccharum*. *Nat. Genet.* **57**:242–253. <https://doi.org/10.1038/s41588-024-02033-w>.
- Zhang, Q., Hua, X., Liu, H., Yuan, Y., Shi, Y., Wang, Z., Zhang, M., Ming, R., and Zhang, J. (2021). Evolutionary expansion and functional divergence of sugar transporters in *Saccharum* (*S. spontaneum* and *S. officinarum*). *Plant J.* **105**:884–906. <https://doi.org/10.1111/tpj.15076>.
- Zhang, Q., Qi, Y., Pan, H., Tang, H., Wang, G., Hua, X., Wang, Y., Lin, L., Li, Z., Li, Y., et al. (2022). Genomic insights into the recent chromosome reduction of autopolyploid sugarcane *Saccharum spontaneum*. *Nat. Genet.* **54**:885–896. <https://doi.org/10.1038/s41588-022-01084-1>.