



A super pan-genome map provides genomic insights into evolution of diploid cotton species

Xueqiang Wang^{1,2,3}, Hejun Lu¹, Yan Zhao⁴, Zhiyuan Zhang², Jun Li^{1,2,3},
Zeyu Dong^{1,2}, Yupeng Hao^{1,2}, Lei Fang¹, Xueying Guan¹, Ting Zhao¹,
Yan Hu¹ and Tianzhen Zhang¹

¹College of Agriculture and Biotechnology, Zhejiang University

²Hainan Institute of Zhejiang University

³Hainan Yazhou Bay Seed Laboratory

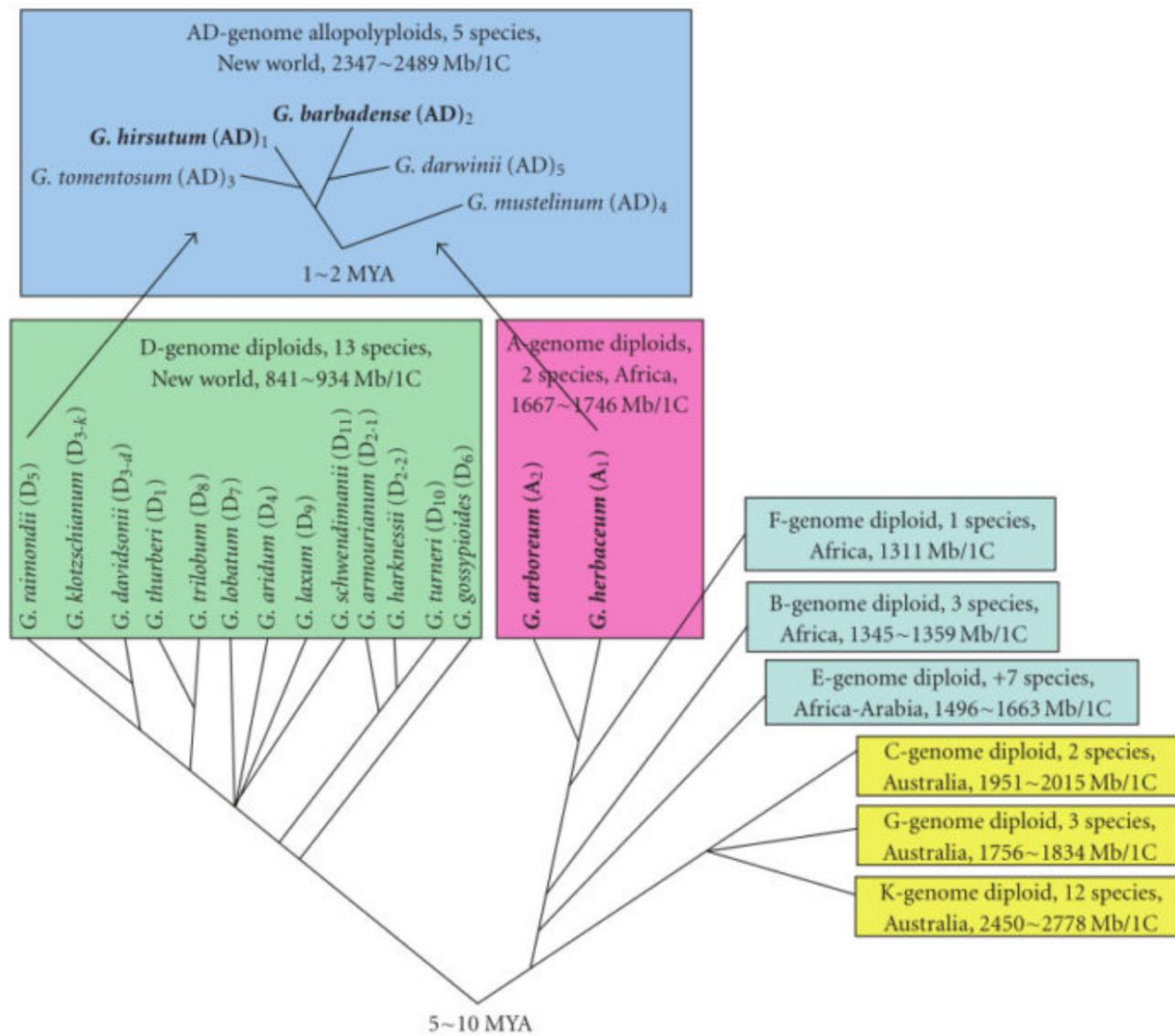
⁴College of Agronomy, Shandong Agricultural University



Xueqiang Wang, Hejun Lu, Yan Zhao, Zhiyuan Zhang, Jun Li, Zeyu Dong, Yupeng Hao, et al. 2024. A super pan-genome map provides genomic insights into evolution of diploid cotton species. *iMetaOmics* 1: e15
<https://doi.org/10.1002/imo2.15>



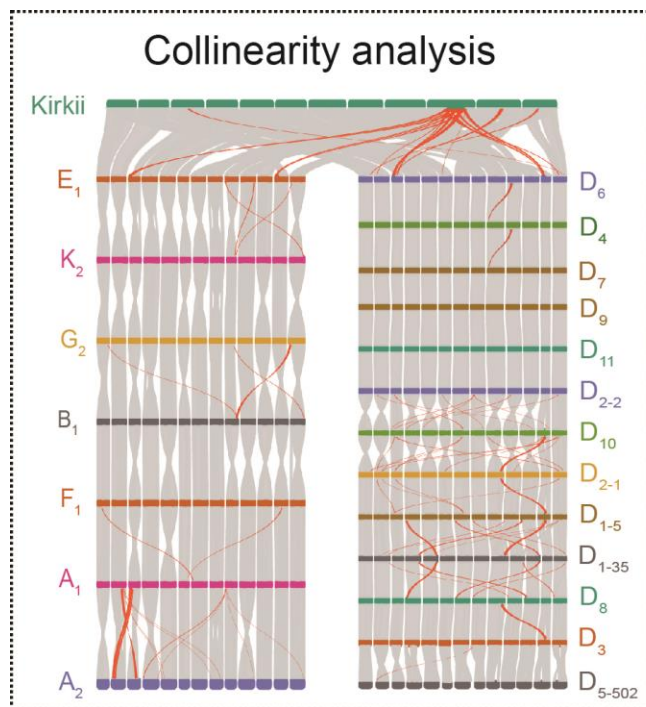
Introduction



- ❑ The pan-genome of all species in the genus *Gossypium* with the D genome has not been constructed;
- ❑ Sequence alignments of foliar nectaries genes from diploid cotton species of the *Gossypium* genus remain incomplete.



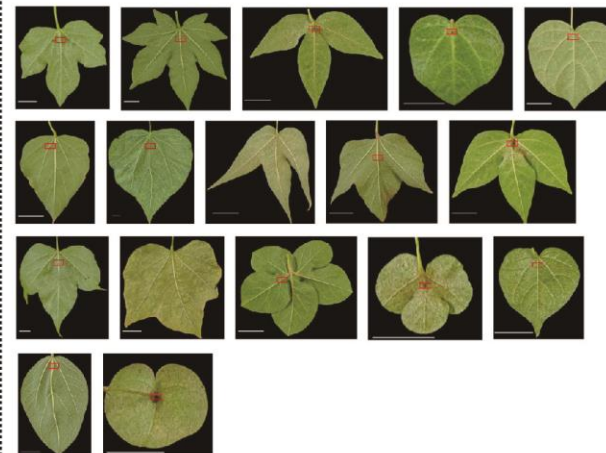
Highlights



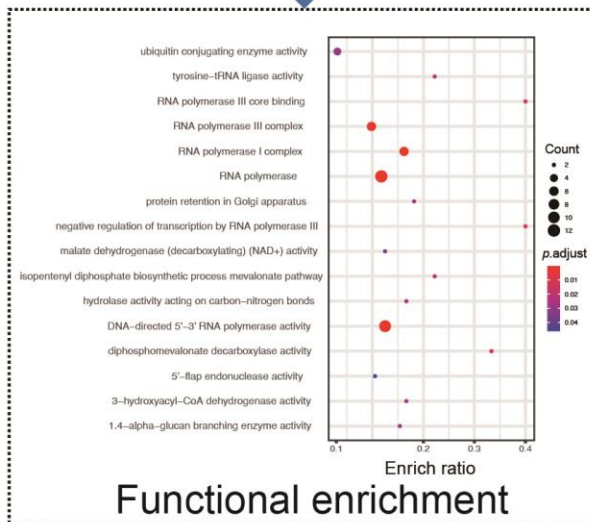
Diploid cotton species



Foliar nectary morphology



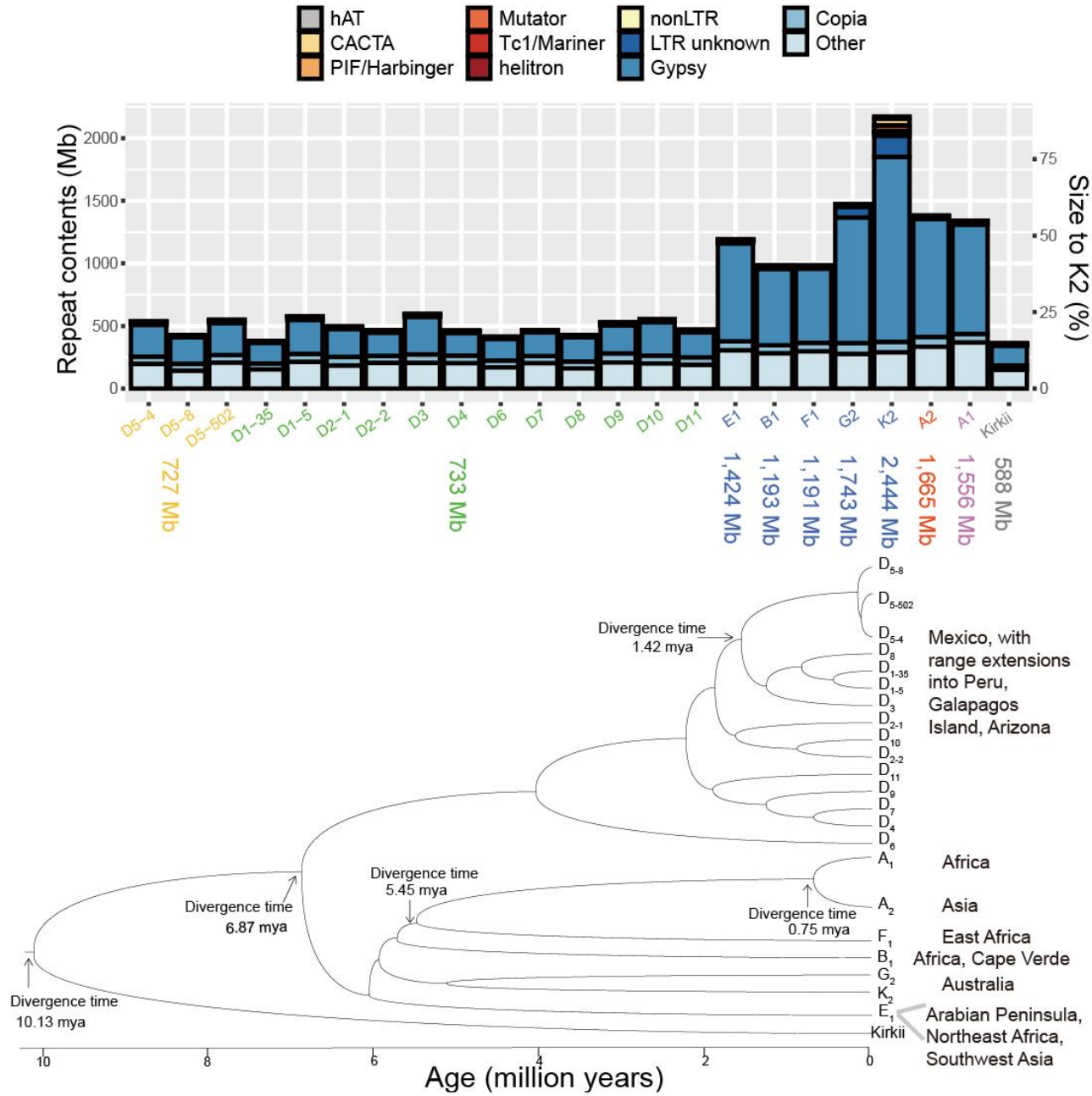
Pan-genome analysis



Structural variations



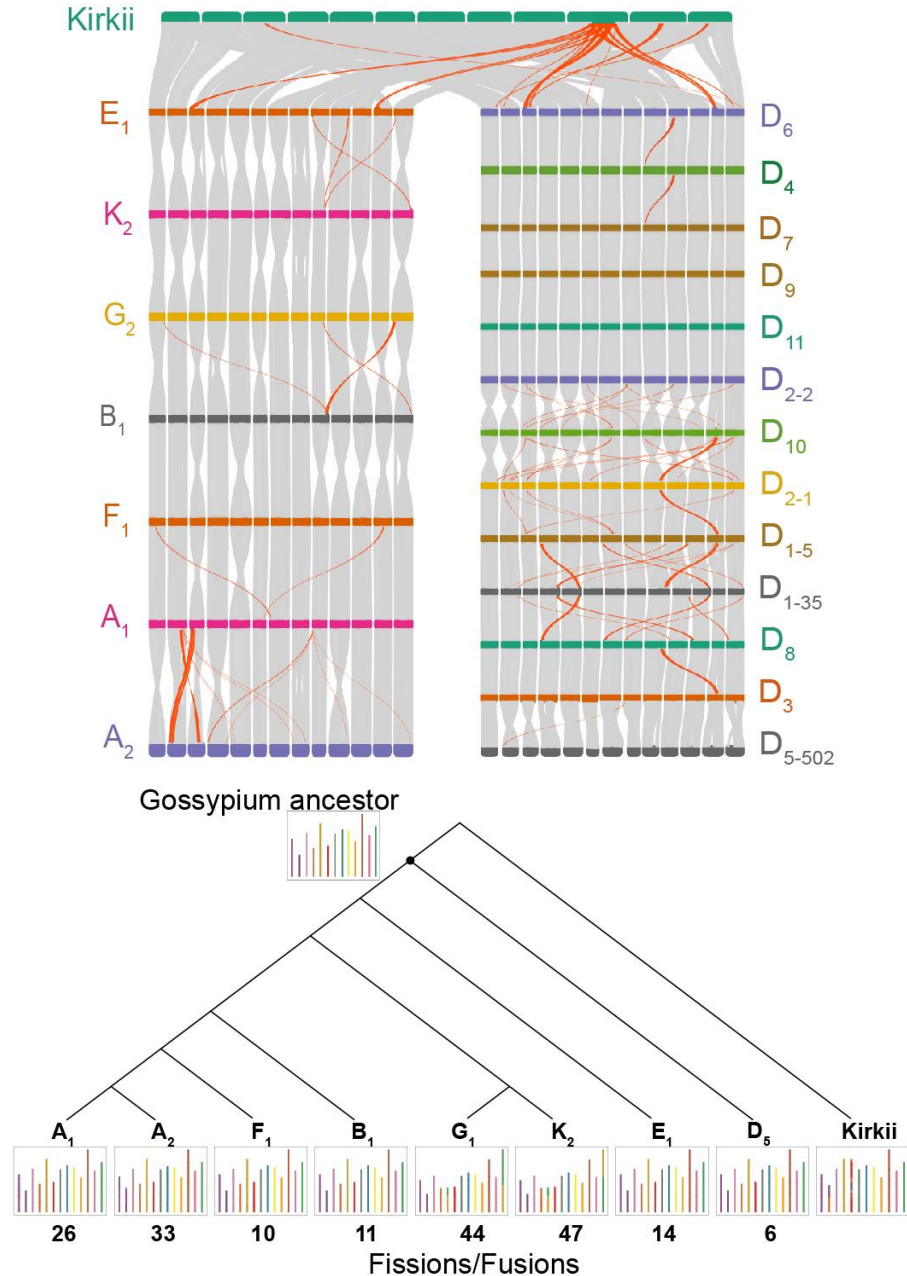
Result



- The most TEs were found in *G. rotundifolium* (K₂), and the fewest in wild relative *Kirkii*;
- There was a significant increase in the length of TEs from the D genome to the G and K genomes, and from the B, E, and F genomes to the A genome;
- There was a significant positive correlation between the proportions of TEs in total sequence per genome and assembly length;
- revealing two distinct clades with D genome diploid cotton species forming one clade;
- The cultivated diploid cotton diverged from the wild diploid cotton species about 5.45 Mya, and the divergence time between diploid cotton species and its wild relative *G. kirkii* was about 10.13 Mya.



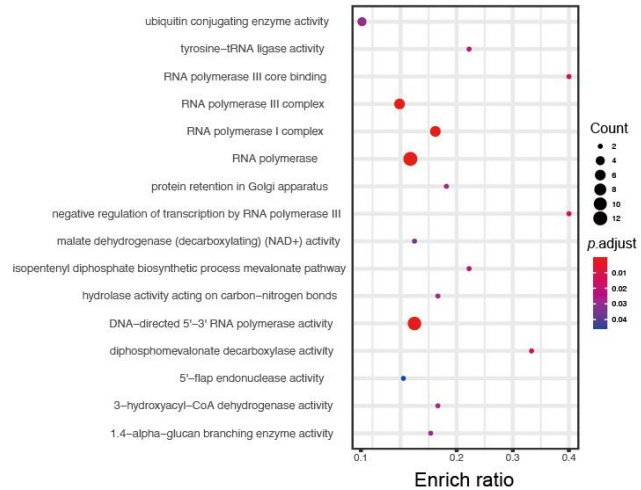
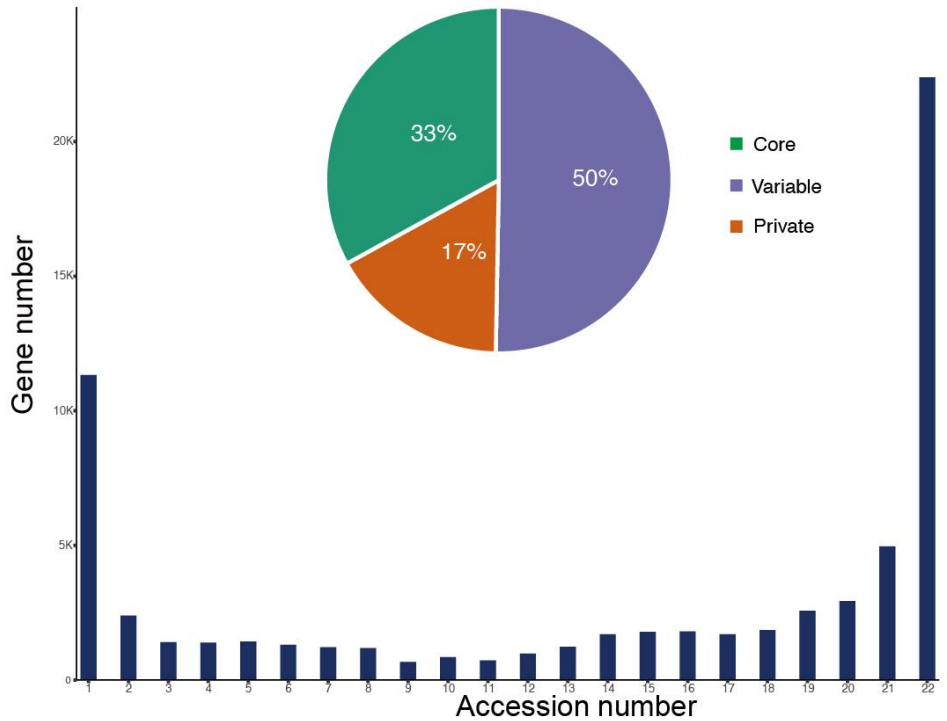
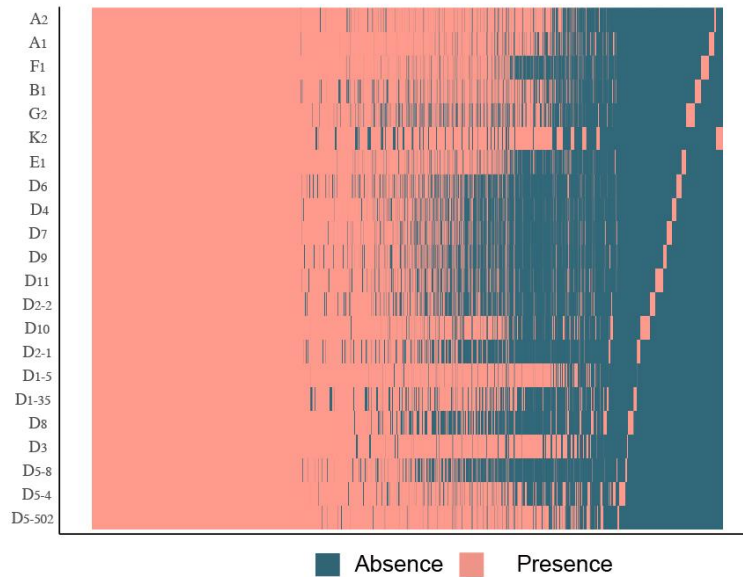
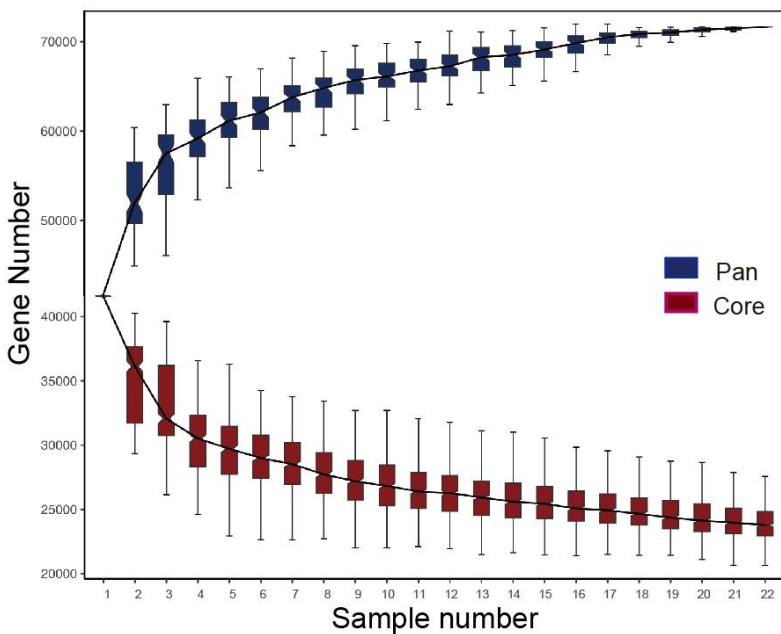
Result



- ❑ Collinearity blocks between assembly genomes were identified, and 13 collinearity blocks between *G. herbaceum* (A₁) and *G. arboreu* (A₂) were disordered on different corresponding genomes, suggesting their importance in diploid cotton evolution;
- ❑ The *Gossypium* ancestor genome was inferred using the GMP model, and chromosome fission and inversions were found to be fundamental forces for speciation;
- ❑ Huge chromosome inversions may drive species formation and diversity.



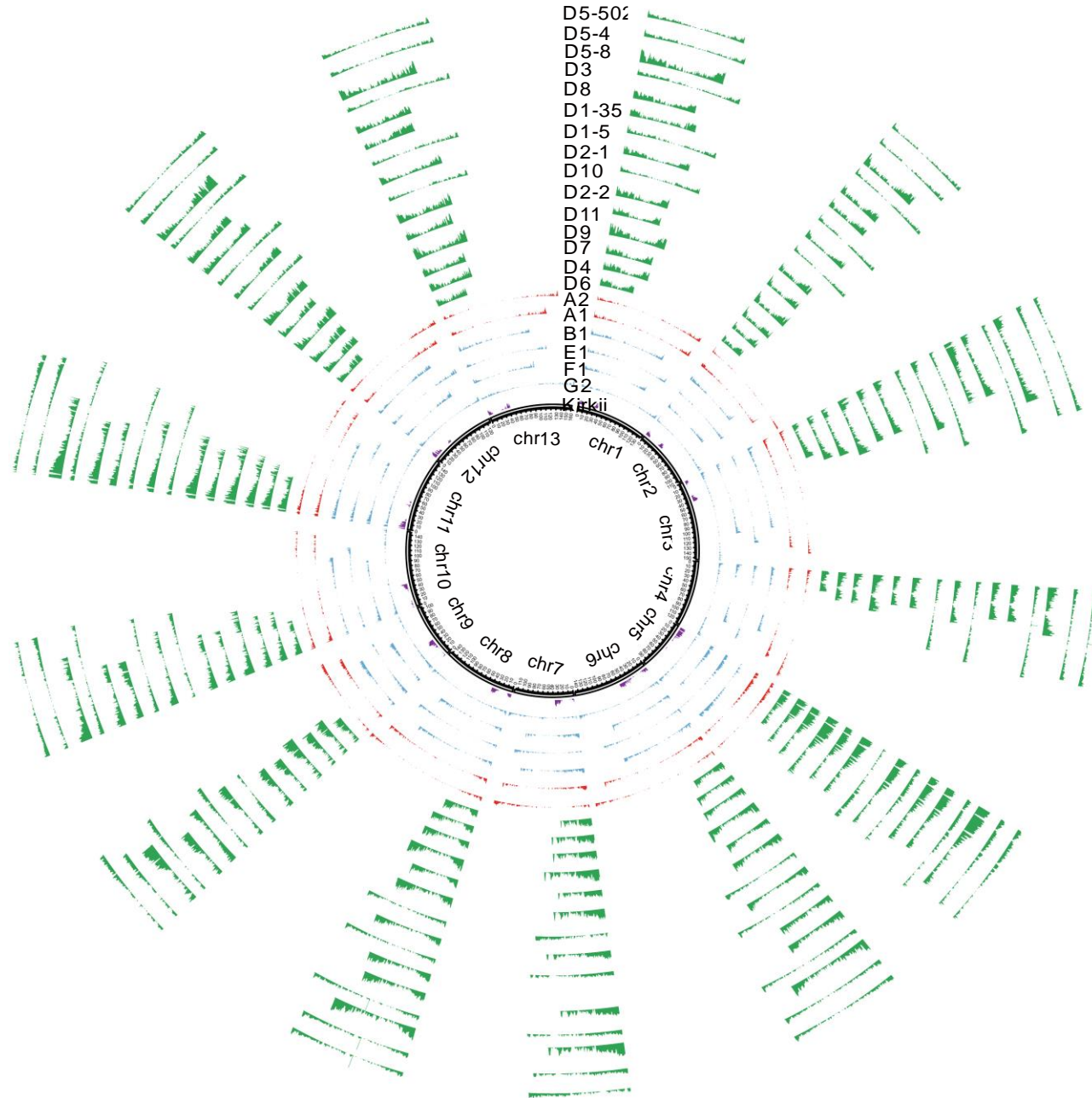
Result



- ❑ The pan-genome contained 67,807 genes, including 22,384 core, 34,093 variable, and 11,330 specific genes;
- ❑ KEGG pathway and GO enrichment analysis of core genes showed terms related to growth and development of cotton;
- ❑ Specific genes were enriched for different terms, revealing variations in characteristics of different cotton species.



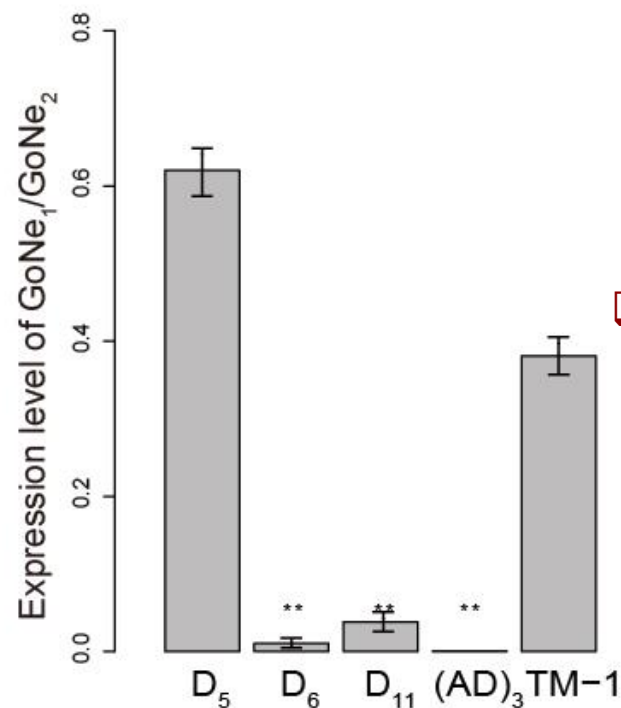
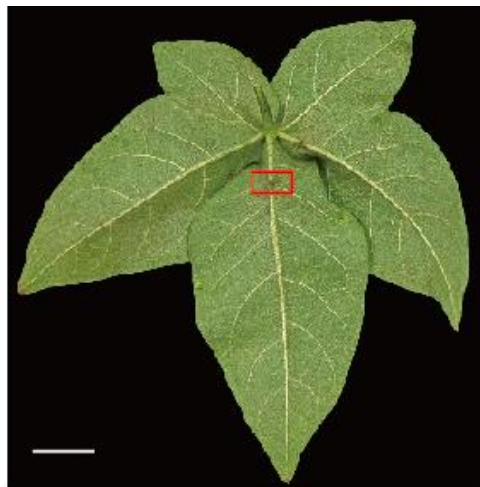
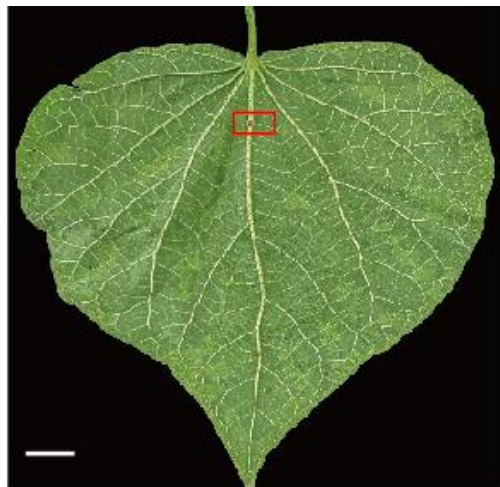
Result



- ❑ Results showed differences in total number and types of SVs across cotton species;
- ❑ Repeat contraction was most common and deletion least common;
- ❑ *G. armourianum* (D₂₋₁) had the greatest number of SVs and *G. kirkii* had the fewest;
- ❑ Unevenly distributed SVs were identified in 321 SV hotspot regions, including 90 genes associated with fiber initiation and/or elongation.



Result



- ❑ No foliar nectary was found in *G. gossypioides* (D₆), *G. schwendimanii* (D₁₁), and *G. tomentosum* ((AD)₃), the allotetraploid cotton species;
- ❑ Comparative genomic analysis of *GoNe* expression revealed no expression in D₆ and D₁₁, suggesting a lack of function of *GoNe* in preventing foliar nectary development in these two wild diploid species;
- ❑ Sequence analysis of the *GoNe* promoter sequences showed a large deletion (444-bp fragment) in the promoter sequences of *GoNe* from D₆ and D₁₁ species compared to other diploid cotton species with foliar nectaries.



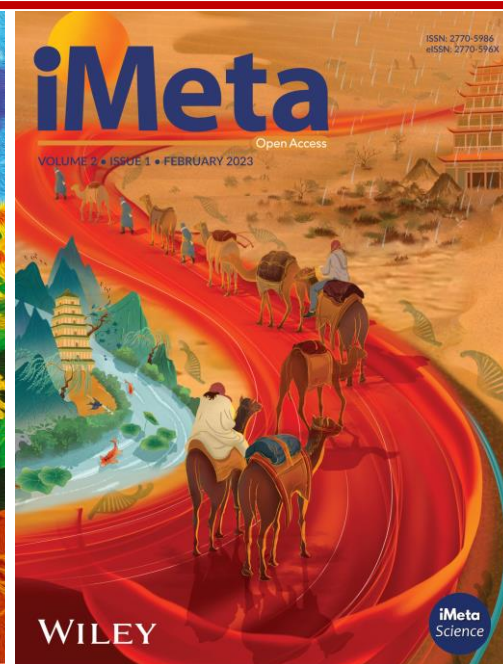
Summary

- ❑ Constructed a high-quality super pan-genome of 22 diploid cotton species and investigated their adaptive evolution;
- ❑ Identified structural variations and hotspot regions associated with fiber initiation and/or elongation;
- ❑ Investigated the absence of foliar nectary in *G. gossypiodes* and *G. schwendimanii*, and identified the deletion in the promoter sequence of *GoNe* as the cause.

Xueqiang Wang, Hejun Lu, Yan Zhao, Zhiyuan Zhang, Jun Li, Zeyu Dong, Yupeng Hao, et al. 2024. A super pan-genome map provides genomic insights into evolution of diploid cotton species. *iMetaOmics* 1: e15
<https://doi.org/10.1002/imo2.15>

iMeta: Integrated meta-omics to change the understanding of the biology and environment


WILEY



“***iMeta***” is a Wiley partner journal launched by iMeta Science Society in 2022, receiving its first impact factor (IF) of **23.7** in 2024, ranking 2/165 in the microbiology field. It aims to publish innovative and high-quality papers with broad and diverse audiences. Its scope is similar to *Nature Biotechnology*, *Nature Microbiology*, and *Cell Host & Microbe*. Its unique features include video abstract, bilingual publication, and social media dissemination, with more than 500,000 followers. It has published 200+ papers and been cited for 4000+ times, and has been indexed by [ESCI/WOS/JCR](#), [PubMed](#), [Google Scholar](#), and [Scopus](#).

“***iMetaOmics***” is a sister journal of “***iMeta***” launched in 2024, with a target IF>10, and its scope is similar to *Microbiome*, *ISME J*, *Nucleic Acids Research*, *Briefings in Bioinformatics*, *Bioinformatics*, etc. All contributes are welcome!

Society: <http://www.imeta.science>
Publisher: <https://wileyonlinelibrary.com/journal/imeta>

 office@imeta.science
imetaomics@imeta.science

 [iMetaScience](#)

Submission: <https://wiley.atyponrex.com/journal/IMT2>
<https://wiley.atyponrex.com/journal/IMO2>

 [Promotion Video](#)

 [iMetaScience](#)